EXPLORING SOCIAL AND COGNITIVE THEORIES OF VOICE PERCEPTION

CHALLENGING ASSUMPTIONS AND EXPLORING NEW APPLICATIONS OF SOCIAL, COGNITIVE, AND EVOLUTIONARY THEORIES OF VOICE PERCEPTION

By JESSICA OSTREGA, B.Sc. (Hons)

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

DESCRIPTIVE NOTE

McMaster University DOCTOR OF PHILOSOPHY (2024) Hamilton, Ontario

(Psychology, Neuroscience & Behaviour)

TITLE: Challenging Assumptions and Exploring New Applications of Social, Cognitive, and Evolutionary Theories of Voice Perception

AUTHOR: Jessica Ostrega, B.Sc. (Hons) (McMaster University)

SUPERVISOR: Dr. David R. Feinberg

PAGES: *xviii* (total preliminary pages), 114 (total pages)

LAY ABSTRACT

This thesis takes a multi-layered approach to examine and challenge existing assumptions regarding the influence of the perception of voice on social judgements. We evaluated whether listening to high- or low-pitched voices during an adaptation condition influenced attractiveness judgements in a similar fashion to what has been found previously in normality judgments. We explored the complexity of social voice judgements based on existing social perception models and first impressions people form from listening to voices. Finally, we addressed how those perceptions might influence person judgements in a novel setting. This work provides a glance into the cognitive, social, and evolutionary complexity of voice impressions but also how those perceptions are used in commonplace interactions.

ABSTRACT

Voice perception is an integral component to social connection and communication. Using the sound of a voice we infer information about a speaker's physical, psychological, and emotional characteristics. These impressions that are formed have the potential to influence behavioural responses to others. This thesis examines some of the fundamental assumptions of voice perception by replicating and extending their findings. In chapter 2, the assumption that exposure to voices alters how attractive voices are was tested. We did not find evidence that increased exposure to high- or low- pitched voices affected attractiveness judgements. Given that exposure to voices did not alter their perceived attractiveness, we were curious to explore if attractiveness judgements were part of first impressions people formed from voices. In chapter 3, we explored what people consciously thought about when listening to voices. We then used machine learning to organize and analyse free form descriptions of participant impressions of voices. A diverse set of topics were used when talking about voices including gender, accent, and social traits. We also confirmed that valence, dominance, and attractiveness were all important social dimensions even when participants were not prompted by researchers to evaluate traits on those domains. We followed these results by testing if the same model of dominance, trust, attractiveness, and competence applied in a practical setting. We had participants judge the voices of doctors and nurses. Low-pitched female voices were perceived as more competent sounding than male voices when they were labelled as belonging to doctors. Low-pitched voices were judged as more dominant regardless of voice sex and profession and high-pitched female voices and low-pitched

male voices were judged as most attractive regardless of profession. We replicated previous findings for attractiveness and dominance perceptions and extended the work by applying it to a novel context. Our findings challenge and expand on existing assumptions of voice perception.

ACKNOWLEDGEMENTS

Firstly, I would like to thank Dr. David Feinberg for the opportunity to work alongside him over the past six years in the Voice Research Lab. The lab has grown and changed in the time I have been part of it, but it was truly and honour to be part of that metamorphosis. David, thank you for seeing my passion for understanding human behaviour, evolution, cognition, and my excitement about voices. Thank you for helping me turn it into this exceptional experience. I am eternally grateful that I was able to explore and discuss such fascinating ideas and concepts with you. I have learned so much from you, and some of the most important lessons I will take with me everywhere I go. You taught me the importance of prioritizing myself and my life beyond academia before anything else. Although I never expected graduate school to be the place I learned about prioritizing a work life balance, I did. You created a lab environment where unlearning, growth, and support was encouraged, I hope that I can one day create a work environment with those same values.

Next, I want to graciously thank Dr. Sandra Monteiro not only for joining my committee at such a critical time, but also for being an incredible mentor when I needed someone to remind me why I started this journey in the first place. Thank you for helping light my passion and drive for science again and showing me all the incredible things, I am capable of. Thank you for sharing your experiences and advice and guiding me to present my best self. Academia can sometimes be a scary place, and no one can dismiss the value of a mentor like you who can help a student see the value in themselves but also

an example of a life that is possible by making the right choices for yourself in academia. I will always cherish our discussions and your leadership.

I have incredible admiration and respect for Dr. Laurel Trainer and the rigor of the science she conducts. I hope to channel the same passion and drive towards excellence as I move forward in my scientific pursuits. I am grateful to have had you on my committee from the beginning of my graduate school journey. Your feedback and suggestions have helped improve my work and improve myself as a researcher. It has been a privilege to be evaluated by you.

Dr. Jennifer Ostovich, it has been a pleasure being your student and working with you over the past 9 years. This thesis would not have been possible without all those key terms that were tenaciously drilled into my head, and it has been and honour learning from you.

To Marie Armstrong, I wish that I could have had more time to learn from you. I will forever remember our late-night lab discussions. The time I had with you as a lab member and friend was invaluable. You fought cancer valiantly and made a permanent mark on the lives of every person who knew you. You were the first person that made me feel ready to take on this challenge. Although it saddens me that you are not here with me as I finish, I know that your soul has more important obligations.

To my book buddy, dear friend and lab mate, Charlene, it was an honour to be there with you as you started and finished your first master's degree. Thank you for getting me out of my lengthy (non-academic) reading slump. I have had such an

incredible experience sharing adventures with you. I am so excited to hear about what you do next and share many more quests with you.

To all other members of the voice lab that have helped me become the researcher and teacher I am today, thank you. Shane, Cibran, Farnia, Justyna, Marquesa, Aaliyah, Laura, Molly, Darren, and all the other amazing lab members you have been a pleasure to work with.

I have probably spent more time in the department than I have spent with my own family in the past five years. You have all helped me grow as a person and to that I am eternally gratefully. To all my fellow past and present graduate students, you are all incredible! I'd especially like to thank Lucas Greville and Portia Kalum for being such amazing mentors, leaders, and friends, Hasan Siddiqui for sharing the lows and highs of the past five years, while keeping our department such an amazing place to exist. Erica, Emily, Brendan, Connie, Sarah, Ben, Hannah, Melise, Ver-Se, Zoe, David, Cindy, Shu, Shreshth, Maheen, Isaac, Yarden, Cam, Janice, the other Lucas, Bethany, Briar, Emily, Allison, Noah, Courtney, Max, Keiko, Rachel, Shane, Rei, Esin, Beth, Devin, Brahm, Jackie, Vini, Sevda, each and every one of you have gotten me here and I hope you know how valuable you are.

To my brother, Martin Ostrega. I know I gave you a hard time throughout your own academic explorations, but I did it with love. Thank you for being there to remind me of how much I have already accomplished and knocking some sense into me when I needed it. I know I still barely know the first thing about salmonid spawning, but I did

read your thesis, so I hope after reading this you finally have an idea of what I have been working on for five years.

To my friends, Shannen, Shannon, Sarah, and Kris, thank you so much for being there to listen to the same rants repeatedly. Thank you for staying with me even when there were months on end where I disappeared from this world. I will never be able to express my gratitude to you all fully.

Harrison, thank you for making sure I ate good food, listened to great music, and drank fantastic wine throughout this time. Without you and your encouragement, I don't think I would have the confidence in myself I do today. Also, thank you for being my translator and advisor as I begin my life beyond graduate school. Who knew a resume could look so pretty?

Mamusia and tatus, there will never be the right words to thank you for all the support and unconditional love you've given me from before I entered this world that has given be everything I needed to succeed today. I hope that this work begins to show how incredibly grateful I am to the sacrifices you've made by coming to Canada so I could have these opportunities. You told me to keep studying, so that's exactly what I did. Thank you for everything and so much more.

Chris, my dear, although you will not receive an honorary PhD for supporting me, thank you for always being there. Your love means the world to me, and I cannot wait to take what I've learned with this experience and start building our life together.

To anyone else who I was not able to formally acknowledge, Thank you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

AI, artificial intelligence

bit, bitrate of a waveform

CAD, Canadian dollar

CI, confidence interval

dB, decibel

d.f., degrees of freedom

F0, fundamental frequency

GBP, Great British pound

Hz, hertz

kHz, kilohertz

LDA, Latent Dirichlet Allocation

N, number of participants

p, p value, probability

PCM, pulse code modulation

RMS, root mean square amplitude

SE, standard error

t, test statistic from a t-test

z, test statistic from a Fisher's z-test

DECLARATION OF ACADEMIC ACHEIVEMENT

This dissertation is organized in the sandwich format, as approved by the McMaster University School of Graduate Studies, comprised of five chapters. Each of the four empirical chapters (Chapters 2-4) is a complete manuscript in preparation for submission.

I am the primary author of each of the three manuscripts. I developed each research question and experiment in this dissertation in consultation with my supervisor, Dr. David Feinberg. I programmed the experiments, collected, and analyzed the data, and authored each chapter. These three manuscripts accurately reflect my doctoral research, and therefore, they comprise the main body of this dissertation. The roles of coauthors for each manuscript are documented below:

CHAPTER 1: General Introduction

I am the sole author of this chapter, which was written in consultation with Dr. David Feinberg.

CHAPTER 2: Adaptation Effects of Voice-Pitch on Attractiveness Judgements

Coauthor Contributions:

Little, A.C. – stimuli, experiment development, manuscript contributions

Feinberg, D.R. – experiment development, data analysis, manuscript contribution

Research Conducted: 2018-2019

CHAPTER 3: Topic Analysis Reveals First Impressions of Voices

Coauthor Contributions:

Feinberg, D.R. – stimuli, experiment development, data analysis, manuscript contribution

Research Conducted: 2020-2022

CHAPTER 4: Are You Listening? Perceptions of Voices Labelled as Doctors and Nurse on Trustworthiness, Competence, Dominance, And Attractiveness Judgements

Coauthor Contributions:

Monteiro, S. – experimental development, manuscript contribution

Feinberg, D.R. – stimuli, experimental development, data analysis, manuscript contribution

Research Conducted: 2023

CHAPTER 5: General Discussion

I am the sole author of this chapter, which was written in consultation with Dr. David Feinberg.

CHAPTER 1: GENERAL INTRODUCTION

Thesis Summary

Voice perception is important for successful social interactions. Hearing people's voices can inform us about their physical, psychological, and mental states. The information we retrieve from voices is used to help us form impressions and ideas about others. This thesis examines different theories about the relationship between voice perception and social judgments. Each empirical study measured different aspects of voice perception and attempted to replicate and extend them. The first theory examined proposes that adaptation to voices changes how attractive they sound. In chapter 2, we tested whether brief exposure to high- and low-pitched voices creates voice attractiveness aftereffects. Exposure to high or low- pitched voices has previously been theorized to increase the perceived normalcy of voices and influence cross-modal representation of speaker gender stereotypes. Here we tested if changing pitch also affects attractiveness judgements in a way that would reflect perceptual adaptation to an attractiveness dimension. We did not find any evidence that increasing exposure to voice pitch affects attractiveness in a way consistent with perceptual adaptation to an attractiveness dimension. We did, however replicate findings that low pitch was attractive in men's voices and high pitch was attractive in women's voices, suggesting the voice stimuli and design are not out of the ordinary.

Having explored how voices are perceived and the dimensions of attractiveness in chapter 2, we shifted our focus to how people naturally think about and describe voices when forming first impressions. In chapter 3, we used machine learning to analyze over

300 people's descriptions of 100 voices saying the word, "Hi". In an unguided analysis, our examination revealed that people thought about a diverse set of topics including age, gender, accent (including, but not limited to language origin and sexual orientation) and social skills. When we guided the analysis to look specifically at valence and dominance, we found that these two dimensions accounted for about one-third of topic keywords. Collectively these results suggest that the dimensions that people use to describe voices are consistent with theoretical models of social perception. Our analysis confirms that valence and dominance are important social dimensions, but also suggests we should focus more research on how perceived accents, age, and emotions affect social voice perception, and cognition.

Given that we found that competence, dominance, trust, and attractiveness were important social dimensions in chapter 4, we examined how these dimensions were evaluated within a healthcare context. In chapter 4, voice stimuli were presented as belonging to medical doctors and nurses. We then explored the relationship between voice perception and judgments of dominance, trust, attractiveness, and competence. Since there are gender stereotypes regarding female doctors and male nurses, we were interested in how perceived occupation interacts with the way voice pitch influences our perceptions of the aforementioned traits. We found that low pitched voices sounded more dominant and competent for both male and female doctors and nurses. Low pitch was more attractive and trustworthy in men's voices and high pitch was more attractive and trustworthy in women's voices.

Low-pitched voices were judged as more dominant regardless of voice sex and profession and high-pitched female voices and low-pitched male voices were judged as most attractive regardless of profession. While our findings suggest that dominance and valence are very important (see chapter 3), we did not find evidence that these dimensions are orthogonal. Although valence and dominance may load onto orthogonal PCA dimensions (McAleer et al., 2014; Oosterhof & Todorov, 2008), voice pitch was correlated with ratings of dominance and competence (dominance dimension) and attractiveness and trust (valence dimension). This finding is consistent with other work that finds that valence and dominance are overlapping dimensions in faces (Sutherland et al., 2013, 2016).

Voice Production

The source filter theory of voice production states that the air moving through the vocal tract (causing vocal-fold vibration) is the sound source, and subpharyngeal vocal tract is an independent filter for the sound (Fant, 1960; Titze, 1994). Voice pitch is the perception of fundamental frequency and corresponding harmonics (Titze, 1994). Voice pitch is usually the lowest frequency of a periodic waveform, but the resonant frequencies (formant frequencies) code for vowels and are also used to estimate vocal-tract size (Fant, 1960; Fitch, 1997).

Voice pitch is the most salient characteristic of voices and the primary correlate of voice attractiveness (Feinberg et al., 2005; Tigue et al., 2012; Zuckerman et al., 1990). Voice pitch is sexually dimorphic in human adults (Titze, 1989). The characteristic of low-pitched voices is strongly associated with higher levels of circulating testosterone

during puberty (Dabbs & Mallinger, 1999; Evans et al., 2008; Feinberg et al., 2005; Harries et al., 1998; Puts et al., 2012). Higher levels of circulating testosterone act to increase the vocal fold size, length and thickness during puberty and deepen a speaker's voice (Harries et al., 1998). Before puberty, sex does not differentiate voice pitch (Abitbol et al., 1999). After puberty, among pre-menopausal same-sex adults, voice pitch is bimodally distributed (Abitbol et al., 1999). After menopause, women's voice pitch can often lower also (Abitbol et al., 1999). Voice pitch is also thought to, but is not associated with body size among same-sex adults (Pisanski et al., 2014).

Voice Categorization and Attractiveness

Voices are thought to be perceived with reference to an internal prototype (Belin et al., 2011). Prototype theory states that a prototype represents an average of the collective experiences an individual has had with a specific stimulus that belong to that category (Rosch, 1973). Therefore, the prototypical voice would be the average voice from all the voices an individual has ever encountered in their lifetime (Belin et al, 2011). The prototypical voice is thought to serve as an anchor in perceptual voice space, when people encounter a novel voice they are thought to do so with reference to their voice prototype.

As such if an individual is presented with a voice that closely resembles their internal prototype they are likely to process the voice faster and with greater ease (Belin et al., 2011; Bruckert et al., 2010; Latinus & Belin, 2011; Winkielman et al., 2006). On the other hand, if the voice does not closely resemble their internal prototype the voice will be processed slower and with some difficulty. It is postulated that voices that are closer to our prototype are thought to be perceived as more attractive because of the ease

with which we process those voices (Winkielman et al., 2006). This is also why average voices are thought to be perceived as more attractive (Belin et al., 2011; Bruckert et al., 2010; Latinus & Belin, 2011). In reference to faces, people have discussed exemplar theory. Using exemplar theory perception is referenced against individuals rather than average representation of social categories. In voices, there is currently a stronger focus on the prototype explanation of voice perception (Belin et al., 2011).

Social Voice Impressions

People are voice perception experts (Latinus & Belin, 2011). We use speech to understand and communicate with one another; this ability makes us a rather unique species. We derive an incredible wealth of socially relevant information simply from listening to a voice. We often infer age, sex, gender, and an array of social traits from listening to a speaker. You can create a detailed picture of the person you are listening to from the sound of their voice (Latinus & Belin, 2011). This picture usually contains some elements of truth, and often has some parts that are completely off the mark. We discuss this below in the *Accuracy* section.

Social Face Space

Since faces and voices are thought to convey similar socially relevant information (Belin et al., 2011), examining social perceptions of faces is important before diving into understanding voice spaces. In the same fashion as understanding voices, people also infer an incredibly rich set of information when they encounter a novel face. Oosterhof & Todorov (2008) asked participants to write everything that came to mind about a person when presented with an image of their face with a neutral expression. Participants then

provided unconstrained impressions of the images of 66 faces. These impressions were then categorized into fourteen personality traits by to independent researchers. The traits were rated by an independent sample of raters on a nine-point scale, where one represented a voice being low on the trait and nine being high on the trait. If the trait was attractiveness, a face that was rated as being low on the trait would be perceived as unattractive whereas a face rated as being high on the trait would be perceived as very attractive. Once these traits were rated, a principal component analysis (PCA) was run to see how these personality characteristic ratings could be summarized. Oosterhof and Todorov (2008) found that the ratings were summarized in a two-dimensional social face space of valence, which highly correlated with trustworthiness judgements and dominance.

To test how well Oosterhof and Todorov's (2008) model applied to ambient images of faces Sutherland et al. (2013) tested their two-dimensional model on 1000 images of faces with a diverse set of facial expressions, glasses, piercing, facial hair, etc. They had their participants rate the faces on traits that were relevant to the current literature. After an obliquely rotated factor analysis, Sutherland et al. (2013) found a novel third factor of youthful attractiveness.

Social Voice Space

McAleer et al. (2014) wanted to test if participants were asked to rate voices on personality characteristics that the social space or dimensions that resulted from those ratings would be consistent with those found in the social face space. To explore the instantaneous impressions that people form from voices they had participants rate 64

recordings of the would “hello” on a set of questions related to dominance and valence: aggressiveness, attractiveness, competence, confidence, dominance, femininity, likeability, masculinity, trustworthiness, and warmth. They found that personality judgements were consistent across listeners, with raters providing similar ratings for the same voices. They were also able to summarize their ratings in a two-dimensional social voice space of valence, that correlated strongly with trustworthiness and likeability, and dominance. McAleer et al. (2014) also found a third dimension of their social voice space, but no rated traits correlated with this third dimension. However, this does suggest that other dimensions of the perceptual voice space might exist. This in conjunction with a third factor being found in other studies that used oblique as opposed to orthogonal factor rotation indicates that a nonlinear analysis might help us find other dimensions in the perceptual voice space.

To examine how people described voices and faces using free form text responses Lavan, (2023) asked participants to provide unconstrained descriptions of six women’s faces and voices to examine if people form categories to both faces and voices that are consistent with previous findings. Participants provided free descriptions of a recording of one of six women’s voices speaking for two minutes, or a muted video of those same six female faces displayed for two minutes. Physical characteristics, such as the age and sex of a person were mentioned earlier than other characteristics. Psychological or personality traits were mentioned most frequently for both voices and faces. Although Lavan (2023) only used six stimuli, these findings showed that we can draw information from voice descriptions. Of particular importance was the finding that personality characteristics

were not the only descriptors people used to form first impressions. Participants not only provided physical traits, but they also assumed people's education, social standing, regional origin, and other socially relevant traits.

Lavan (2023) also continues to take the approach that Oosterhof & Todorov (2008) used when examining face impressions. Using unconstrained responses removes a level of researcher decision making, decisions that can influence the characterization of dimensions (Mondloch et al., 2022). Moving away from the use of rating scales prevents a reduction of potential characteristics described by participants and helps control for stereotype and biases that might unintendedly be imposed by such constraints (Mondloch et al., 2022).

Personality Judgements

The perceptions of an individual's personality are strongly influenced by their voice. Men and women with lower pitched voices are more likely to be chosen for leadership positions than those with higher pitched voices (Anderson & Klofstad, 2012; Klofstad et al., 2012; Tigue et al., 2012). Lower pitched voices of leaders were also perceived as more competent and stronger than leaders with higher pitched voices (Klofstad et al., 2015). Leaders in their 40s and 50s are preferred over those in their 30s, 60s and 70s (Klofstad et al., 2015). In electoral decision making when facing male opponents, candidates with lower pitched voices receive more votes than those with higher pitched voices (Klofstad, 2016). Conversely participants with higher pitched voices received more votes when facing female opponents, this was especially true when candidates were male (Klofstad, 2016). Voice pitch has also been found to negatively

correlate with people's self-reported sociosexuality, dominance, and extraversion (Stern et al., 2021). Even children rate low-pitched men's voices as more competent than high-pitched voices for stereotypically male occupations (Cartei et al., 2021). Perceptions of people's trustworthiness and dominance are strongly driven by their facial (Krugar, 2006; O'Connor & Feinberg, 2012; Smith et al., 2009; Stirrat & Perrett, 2010) and vocal features (O'Connor & Barclay, 2017; Rezlescu et al., 2015), these traits can also influence perceptions of other important characteristics (Lavan, 2023; McAleer et al., 2014; Oosterhof & Todorov, 2008; Scherer, 1972).

Trustworthiness Judgements

There are mixed results on how voice pitch influences perceptions of trustworthiness. Some work finds that people tend to trust higher pitched male voices more than lower pitched male voice both generally (McAleer, Todorov, & Belin, 2014; O'Connor & Barclay, 2017) and in financial contexts (Montano, Tigue, Isenstein, Barclay, & Feinberg, 2017; O'Connor & Barclay, 2017). Other findings indicate that lower pitched voices are generally perceived are more trustworthy than higher pitched voices in both male (Oleszkiewicz, Pisanski, Lachowicz-Tabaczek, & Sorokowska, 2017; Tigue et al., 2012) and female speakers (Klofstad, Anderson, & Peters, 2012). While other studies have not detected any significant effects of voice pitch on perceptions of trustworthiness (Klofstad, Anderson, & Peters, 2012; Vukovic et al., 2011). Although low-pitched male voices and high-pitched female voices correlate with attractiveness judgments (reviewed in Feinberg, 2008; Puts, Jones, & DeBruine, 2012), people generally perceive lower pitched male voices and higher pitched female voices as sounding untrustworthy in

romantic contexts (O'Connor & Feinberg, 2012; O'Connor, Pisanski, Tigue, Fraccaro, & Feinberg, 2014; O'Connor, Re, & Feinberg, 2011). Therefore, the influence of voice pitch on trustworthiness perceptions is equivocal.

Accuracy of Impressions

Agreement exists between personality judgements people make from images of faces and the sounds of voices (McAleer et al., 2014; Oosterhof & Todorov, 2008; Penton-Voak et al., 2006; Todorov et al., 2015; Yovel & Belin, 2013; Zebrowitz & Montepare, 2008). This similarity in judgements does not necessarily represent their accuracy. It is important to note that studies that find a relationship between accuracy and impressions of personality, political orientation, sexual orientation, and criminal behaviour are often rooted in societal biases (Stern et al., 2021). When properties of gender, ethnicity, and age are controlled, the accuracy of the impressions people make become no greater than chance (Todorov et al., 2015). There is evidence that attractiveness judgements of faces and voices are correlated (Feinberg, 2008; Hughes & Miller, 2016; Wells et al., 2013) and that people have the ability to accurately match and individual face to their voice at better than chance accuracy (Kamachi et al., 2003; Mavica & Barenholtz, 2013).

Averageness and Attractiveness

There are two primary hypotheses associating averageness and attractiveness. The first is the *averageness is attractive* hypothesis (Langlois & Roggman, 1990). This hypothesis is supported if the magnitude of distance, but not the direction, from average influences attractiveness. In other words, increasing the distance from the average should

always decrease attractiveness, regardless of the direction. If voice pitch causes voice attractiveness aftereffects, hearing high-pitched voices or low-pitched voices makes them both sound more average and therefore more attractive. It would not matter if someone preferred men or women with high or low voice pitch, hearing either category more has the same effect of increasing averageness and attractiveness.

Support for the relationship between averageness and attractiveness is equivocal. Bruckert et al. (2010) tested whether morphing voices together increases the attractiveness ratings of those voices. Using a sample of 32 male and 32 female adult voices they created average composites (Bruckert et al., 2010). The 32 voices were randomly paired with one another generating sixteen 2-voice composites. This process was repeated at subsequent degrees of averaging to yield eight 4-voice composites, four 8-voice composites, two 16-voice composites, and a single 32-voice composite for each gender. Participants provided attractiveness ratings for natural and composites of female voices first and then male voices. Bruckert et al. (2010) found that in a small sample of 25 participants (13 female, 12 male), the number of voices in the composites positively predicted attractiveness ratings. However, when we attempted to replicate this finding, we found that averaging voices together did decrease distinctiveness, but it did not have any effect on averageness (Ostrega et al., 2024).

An alternative explanation of how averageness affects the attractiveness of voices and faces is the contrast hypothesis (DeBruine et al., 2007). This suggests that attractiveness is a vector in multidimensional voice space and consequently both the magnitude *and* direction away from average play a role in attractiveness. Support for this

idea comes from many experiments showing that people are attracted to low-pitched voices in men and high-pitched voices in women (reviewed in Pisanski & Feinberg, 2018). The idea is that high-pitched voices among women and low-pitched voices in men are both attractive and can be equally distant from average in opposite directions.

Suppose low voice pitch is attractive in men (Pisanski & Feinberg, 2018), and the direction from average matters. If we expose people to low voice pitch, an aftereffect will make low pitched men's voices sound higher in pitch, and therefore less attractive. The opposite would be true for high pitched voices. Exposure to high-pitched voices would make them sound lower in pitch and more attractive. In both cases, the perception of pitch moved towards average, but would have different effects on attractiveness based on the starting point. For women's voices, since high pitch is often found more attractive (Borkowska & Pawlowski, 2011; Feinberg et al., 2005; Re et al., 2012; Röder et al., 2013), we would make the opposite predictions with reference to the effect of direction from average on attractiveness.

Mere Exposure

The more people are exposed to specific stimulus, including voices, the more attractive or likeable that stimuli become with future exposure (Reis et al., 2011; Zajonc, 1968). This mere exposure effect is likely a result of liking things that are familiar to us. This also begins to relate back to how we categorize objects and sounds. If the prototypes or exemplars in our mind easily present us with information that allow us to categorize a novel stimulus quickly because we have been exposed to it many times and can retrieve the stimuli from memory easily, we might attribute that familiarity with liking it. If people

are presented with the same voice repeatedly they are more likely to recognize it and attribute the ease of processing the voice to liking rather than familiarity (Zajonc, 1968).

The difference between mere exposure and adaptation is that perceptual adaptation is a mechanism by which specific neural responses selectively increase or decrease after prolonged or repetitive exposure to stimuli (Clifford & Rhodes, 2005). This neural response is accompanied by a perceptual aftereffect that presents itself for a period immediately after exposure to that stimulus (Clifford & Rhodes, 2005; Thompson & Burr, 2009). Perceptual adaptation effects will only last for a finite length of time after exposure to stimuli (Thompson & Burr, 2009), while the mere exposure effect is influenced by being shown the same stimuli multiple times and can alter an internal categorization of a stimuli as familiar or unfamiliar (Zajonc, 1968). Mere exposure is thought to occur at a higher level of processing. Both share the common thread that familiarity is attractive in this instance, but changes in aesthetic preferences are not a mechanistic part adaptation, but rather a downstream by-product.

The Current Dissertation

In chapter 2 of this dissertation, we examined whether exposure to voice pitch, the primary correlate of voice attractiveness (Feinberg et al., 2005; Tigue et al., 2012; Zuckerman et al., 1990), produces attractiveness aftereffects. We did this by testing people's preferences for high- and low-pitched voices, presenting participants with a set of the same voices manipulated to have higher or lower than average voice pitch, and measuring the change in preferences for attractive voice qualities. Although the same procedure affected normality judgements in prior work (Little et al., 2013), and we

replicated findings that high pitch in women and low pitch in men were attractive (Apicella & Feinberg, 2009; Borkowska & Pawlowski, 2011; Cussigh et al., 2020; Feinberg et al., 2005, 2008, 2011; Jones et al., 2010; Re et al., 2012; Schild et al., 2020; Simmons et al., 2011), exposing people to high or low pitch did not produce attractiveness aftereffects.

In chapter 3, we analyzed the categories people use to describe unfamiliar people's voices saying "Hi". Participants listened to 100 voice (50 female and 50 male) and described in as much detail as possible their impression of each speaker in a text box. We analyzed the data using Latent Dirichlet Allocation, a popular topic analysis algorithm. We found that the most important topics people described voices with were (in order): age, valence, gender and sexual orientation, linguistics and accents, dominance, education and professional context, media and entertainment, and physical appearance. These results were consistent with previous findings in face and voice judgements (Lavan, 2023; McAleer et al., 2014; Oosterhof & Todorov, 2008; Shiramizu et al., 2022; Sutherland et al., 2013) and support the idea that the voice conveys a rich set of information. These results begin to unravel the complexities of the social impressions people form of others, even just from hearing them say the word "hi".

Given the support for the dominance and valence theory of voice perception (McAleer et al., 2014; Oosterhof & Todorov, 2008; Shiramizu et al., 2022) in chapter 4, we investigated the role of voice pitch in attractiveness, competence, dominance, and trust in a medical scenario. With the increase in care provision of doctors and nurses over telephone and online communication, perception of healthcare workers can potentially

impact patient adherence to medical care and their likelihood of seeking care. We found that when labelled as belonging to doctors, low-pitched female voices were perceived as more competent sounding than male voices. Low-pitched voices were judged as more dominant regardless of voice sex and profession and high-pitched female voices and low-pitched male voices were judged as most attractive regardless of profession. Our finding highlight how perceptual psychology might be used to understand judgement in fields like healthcare.

While we found that people preferred low pitch in men's voices, and high pitch in women's voices, exposure to voices varying on pitch, used here as a proxy to the attractiveness dimension, did not create aftereffects consistent with either hypothesis. The effects of pitch on attractiveness remain stable even after brief exposure to groups of voices varying in pitch.

Voices and faces both carry a rich set of socially relevant information, previous work has found that dominance and valence dimensions are important in judgements of people's faces. We also found that valence and dominance attributes were important in people judgements of voices among other characteristics (Oosterhof & Todorov, 2008; Sutherland et al., 2013). Moving away from constraining participant responses using rating scales, this dissertation begins to examine and present the richness of importation that can be gathered from unconstrained and free form text responses of voice perception (Mondloch et al., 2022). This further encourages the examination of person perception using unrestricted descriptions. Both methods converge on similar viewpoints, and bring

their strengths in supporting the idea that voice conveys information about accent, age, gender, dominance, valence, and much more.

Finally, this piece of work adds to the existing literature on how person perceptions, even just from brief utterances of a person's voice can influence everyday interactions. Taking what we have understood about the complexities of judgements made from voices and applying them to how trust, competence, dominance, and attractiveness are perceived in doctors and nurses in potential patients. With the goal of encouraging care providers to be mindful of how some person characteristics such voices might be perceived by those receiving care.

REFERENCES

Abitbol, J., Abitbol, P., & Abitbol, B. (1999). Sex hormones and the female voice. *Journal of Voice*, *13*(3), 424–446. https://doi.org/10.1016/S0892-1997(99)80048-4

Anderson, R. C., & Klofstad, C. A. (2012). Preference for Leaders with Masculine Voices Holds in the Case of Feminine Leadership Roles. *PLoS ONE*, *7*(12), e51216. https://doi.org/10.1371/journal.pone.0051216

Apicella, C. L., & Feinberg, D. R. (2009). Voice pitch alters mate-choice-relevant perception in hunter–gatherers. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1659), 1077–1082. https://doi.org/10.1098/rspb.2008.1542

Belin, P., Bestelmeyer, P. E. G., Latinus, M., & Watson, R. (2011). Understanding Voice Perception: Understanding voice perception. *British Journal of Psychology*, *102*(4), 711–725. https://doi.org/10.1111/j.2044-8295.2011.02041.x

Borkowska, B., & Pawlowski, B. (2011). Female voice frequency in the context of dominance and attractiveness perception. *Animal Behaviour*, *82*(1), 55–59. https://doi.org/10.1016/j.anbehav.2011.03.024

Bruckert, L., Bestelmeyer, P., Latinus, M., Rouger, J., Charest, I., Rousselet, G. A., Kawahara, H., & Belin, P. (2010). Vocal Attractiveness Increases by Averaging. *Current Biology*, *20*(2), 116–120. https://doi.org/10.1016/j.cub.2009.11.034

Cartei, V., Oakhill, J., Garnham, A., Banerjee, R., & Reby, D. (2021). Voice Cues Influence Children's Assessment of Adults' Occupational Competence. *Journal of Nonverbal Behavior*, *45*(2), 281–296. https://doi.org/10.1007/s10919-020-00354-y

Clifford, C. W. G., & Rhodes, G. (2005). *Fitting the Mind to the World: Adaptation and After-Effects in High-Level Vision*. OUP Oxford.

Cussigh, G., Ballester-Arnal, R., Gil-Llario, M. D., Giménez-García, C., & Castro-Calvo, J. (2020). Fundamental frequency of the female's voice: A cross-country empirical study on its influence on social and sexual selection. *Personality and Individual Differences*, *160*, 109937. https://doi.org/10.1016/j.paid.2020.109937

Dabbs, J. M., & Mallinger, A. (1999). High testosterone levels predict low voice pitch among men. *Personality and Individual Differences*, *27*(4), 801–804. https://doi.org/10.1016/S0191-8869(98)00272-4

DeBruine, L. M., Jones, B. C., Unger, L., Little, A. C., & Feinberg, D. R. (2007). Dissociating averageness and attractiveness: Attractive faces are not always average. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(6), 1420–1430. https://doi.org/10.1037/0096-1523.33.6.1420

Evans, S., Neave, N., Wakelin, D., & Hamilton, C. (2008). The relationship between testosterone and vocal frequencies in human males. *Physiology & Behavior*, *93*(4–5), 783–788. https://doi.org/10.1016/j.physbeh.2007.11.033

Fant, F. (1960). *Acoustic theory of speech prodction*. The Hague, Netherlands: Mouton.

Feinberg, D. R. (2008). Are human faces and voices ornaments signaling common underlying cues to mate value? *Evolutionary Anthropology: Issues, News, and Reviews*, *17*(2), 112–118. https://doi.org/10.1002/evan.20166

Feinberg, D. R., DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2008). The Role of

Femininity and Averageness of Voice Pitch in Aesthetic Judgments of Women's Voices. *Perception*, *37*(4), 615–623. https://doi.org/10.1068/p5514

Feinberg, D. R., Jones, B. C., DeBruine, L. M., O'Connor, J. J. M., Tigue, C. C., & Borak, D. J. (2011). Integrating fundamental and formant frequencies in women's preferences for men's voices. *Behavioral Ecology*, *22*(6), 1320–1325. https://doi.org/10.1093/beheco/arr134

Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour*, *69*(3), 561–568. https://doi.org/10.1016/j.anbehav.2004.06.012

Fitch, W. T. (1997). *Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques*.

Harries, M., Hawkins, S., Hacking, J., & Hughes, I. (1998). Changes in the male voice at puberty: Vocal fold length and its relationship to the fundamental frequency of the voice. *The Journal of Laryngology & Otology*, *112*(5), 451–454. https://doi.org/10.1017/S0022215100140757

Hughes, S. M., & Miller, N. E. (2016). What sounds beautiful looks beautiful stereotype: The matching of attractiveness of voices and faces. *Journal of Social and Personal Relationships*, *33*(7), 984–996. https://doi.org/10.1177/0265407515612445

Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C., & Vukovic, J. (2010). A

domain-specific opposite-sex bias in human preferences for manipulated voice pitch. *Animal Behaviour*, *79*(1), 57–62. https://doi.org/10.1016/j.anbehav.2009.10.003

Kamachi, M., Hill, H., Lander, K., & Vatikiotis-Bateson, E. (2003). 'Putting the Face to the Voice': Matching Identity across Modality. *Current Biology*.

Klofstad, C. A. (2016). Candidate Voice Pitch Influences Election Outcomes: Voice Pitch Influences Voters. *Political Psychology*, *37*(5), 725–738. https://doi.org/10.1111/pops.12280

Klofstad, C. A., Anderson, R. C., & Nowicki, S. (2015). Perceptions of Competence, Strength, and Age Influence Voters to Select Leaders with Lower-Pitched Voices. *PLOS ONE*, *10*(8), e0133779. https://doi.org/10.1371/journal.pone.0133779

Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: Voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1738), 2698–2704. https://doi.org/10.1098/rspb.2012.0311

Kruger, D. J. (2006). Male facial masculinity influences attributions of personality and reproductive strategy. Personal Relationships, 13, 451–463. http://dx.doi.org/10.1111/j. 1475-6811.2006.00129.x.

Langlois, J. H., & Roggman, L. A. (1990). Attractive Faces Are Only Average. *Psychological Science*, *1*(2), 115–121. https://doi.org/10.1111/j.1467-9280.1990.tb00079.x

Latinus, M., & Belin, P. (2011). Anti-Voice Adaptation Suggests Prototype-Based Coding

of Voice Identity. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00175

Lavan, N. (2023). How do we describe other people from voices and faces? *Cognition*, *230*, 105253. https://doi.org/10.1016/j.cognition.2022.105253

Little, A. C., Feinberg, D. R., DeBruine, L. M., & Jones, B. C. (2013). Adaptation to Faces and Voices: Unimodal, Cross-Modal, and Sex-Specific Effects. *Psychological Science*, *24*(11), 2297–2305. https://doi.org/10.1177/0956797613493293

Mavica, L. W., & Barenholtz, E. (2013). Matching Voice and Face Identity From Static Images. *Journal of Experimental Psychology-Human Perception and Performance*, *39*(2), 307–312. https://doi.org/10.1037/a0030945

McAleer, P., Todorov, A., & Belin, P. (2014). How Do You Say 'Hello'? Personality Impressions from Brief Novel Voices. *PLoS ONE*, *9*(3), e90779. https://doi.org/10.1371/journal.pone.0090779

Mondloch, C. J., Twele, A. C., & Thierry, S. M. (2022). We need to move beyond rating scales, white faces and adult perceivers: Invited Commentary on Sutherland & Young (2022), understanding trait impressions from faces. *British Journal of Psychology*, bjop.12619. https://doi.org/10.1111/bjop.12619

Montano, K. J., Tigue, C. C., Isenstein, S. G. E., Barclay, P., & Feinberg, D. R. (2017). Men's voice pitch influences women's trusting behaviour. Evolution and Human Behavior. http://dx.doi.org/10.101/j.evolhumbehav.2016.10.010

O'Connor, J. J. M., & Barclay, P. (2017). The influence of voice pitch on perceptions of

trustworthiness across social contexts. Evolution and Human Behavior, 38, 506–512. https://doi.org/10.1016/j.evolhumbehav.2017.03.001

O'Connor, J. J. M., & Feinberg, D. R. (2012). The influence of facial masculinity and voice pitch on jealousy and perceptions of intrasexual rivalry. Personality and Individual Differences, 52, 369–373. http://dx.doi.org/10.1016/j.paid.2011.10.036.

O'Connor, J. J. M., Pisanski, K., Tigue, C. C., Fraccaro, P. J., & Feinberg, D. R. (2014). Perceptions of infidelity risk predict women's preferences for low male voice pitch in short-term over long-term relationship contexts. *Personality and Individual Differences*, *56*, 73–77. https://doi.org/10.1016/j.paid.2013.08.029

O'Connor, J. J. M., Re, D. E., & Feinberg, D. R. (2011). Voice pitch influences perceptions of sexual infidelity. Evolutionary Psychology, 9, 64–78. http://dx.doi.org/10.1177/ 147470491100900109.

Oleszkiewicz, A., Pisanski, K., Lachowicz-Tabaczek, K., & Sorokowska, A. (2017). Voice-based assessments of trustworthiness, competence, and warmth in blind and sighted adults. *Psychonomic Bulletin & Review*, *24*(3), 856–862. https://doi.org/10.3758/s13423-016-1146-y

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Ostrega, J., Shiramizu, V., Lee, A. J., Jones, B. C., & Feinberg, D. R. (2024). No evidence that averaging voices influences attractiveness. *Scientific Reports*, *14*(1), 10488.

Penton-Voak, I. S., Pound, N., Little, A. C., & Perrett, D. I. (2006). Personality judgments

from natural and composite facial images: More evidence for a “kernel of truth” in social perception. *Social Cognition*, *24*(5), 607–640.

Pisanski, K., & Feinberg, D. R. (2018). *Oxford handbook of voice perception.*

Pisanski, K., Fraccaro, P. J., Tigue, C. C., O’Connor, J. J. M., Röder, S., Andrews, P. W., Fink, B., DeBruine, L. M., Jones, B. C., & Feinberg, D. R. (2014). Vocal indicators of body size in men and women: A meta-analysis. *Animal Behaviour*, *95*, 89–99. https://doi.org/10.1016/j.anbehav.2014.06.011

Puts, D. A., Apicella, C. L., & Cárdenas, R. A. (2012). Masculine voices signal men’s threat potential in forager and industrial societies. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1728), 601–609. https://doi.org/10.1098/rspb.2011.0829

Re, D. E., O’Connor, J. J. M., Bennett, P. J., & Feinberg, D. R. (2012). Preferences for Very Low and Very High Voice Pitch in Humans. *PLoS ONE*, *7*(3), e32719. https://doi.org/10.1371/journal.pone.0032719

Reis, H. T., Maniaci, M. R., Caprariello, P. A., Eastwick, P. W., & Finkel, E. J. (2011). Familiarity does indeed promote attraction in live interaction. *Journal of Personality and Social Psychology*, *101*(3), 557–570. https://doi.org/10.1037/a0022885

Rezlescu, C., Penton, T., Walsh, V., Tsujimura, H., Scott, S. K., & Banissy, M. J. (2015). Dominant voices and attractive faces: The contribution of visual and auditory information to integrated person impressions. Journal of Nonverbal Behavior, 39, 355–370. https://doi.org/10.1007/ s10919-015-0214-8

Röder, S., Fink, B., & Jones, B. C. (2013). Facial, Olfactory, and Vocal Cues to Female Reproductive Value. *Evolutionary Psychology*, *11*(2), 147470491301100. https://doi.org/10.1177/147470491301100209

Rosch, E. (1973). Prototype theory. *Cognitive Development and the Acquisition of Language*, 111–144.

Scherer, K. R. (1972). Judging personality from voice: A cross-cultural approach to an old issue in interpersonal perception 1. *Journal of Personality*, *40*(2), 191–210.

Schild, C., Aung, T., Kordsmeyer, T. L., Cardenas, R. A., Puts, D. A., & Penke, L. (2020). Linking human male vocal parameters to perceptions, body morphology, strength and hormonal profiles in contexts of sexual selection. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-77940-z

Shiramizu, V. K. M., Lee, A. J., Altenburg, D., Feinberg, D. R., & Jones, B. C. (2022). The role of valence, dominance, and pitch in perceptions of artificial intelligence (AI) conversational agents' voices. *Scientific Reports*, *12*(1), 22479. https://doi.org/10.1038/s41598-022-27124-8

Simmons, L. W., Peters, M., & Rhodes, G. (2011). Low Pitched Voices Are Perceived as Masculine and Attractive but Do They Predict Semen Quality in Men? *PLoS ONE*, *6*(12), e29271. https://doi.org/10.1371/journal.pone.0029271

Smith, F. G., Jones, B. C., Welling, L. L. W., Little, A. C., Vukovic, J., Main, J. C., & DeBruine, L. M. (2009). Waist-hip ratio predicts women's preferences for masculine male faces, but not perceptions of men's trustworthiness. Personality

and Individual Differences, 47, 476–480. https://doi.org/10.1016/j.paid.2009.04.022

Stern, J., Schild, C., Jones, B. C., DeBruine, L. M., Hahn, A., Puts, D. A., Zettler, I., Kordsmeyer, T. L., Feinberg, D., Zamfir, D., Penke, L., & Arslan, R. C. (2021). Do voices carry valid information about a speaker's personality? *Journal of Research in Personality*, *92*, 104092. https://doi.org/10.1016/j.jrp.2021.104092

Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. Psychological Science, 21, 349–354. https://doi.org/10.1177/09567976103 62647

Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Michael Burt, D., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition*, *127*(1), 105–118. https://doi.org/10.1016/j.cognition.2012.12.001

Sutherland, C. A. M., Oldmeadow, J. A., & Young, A. W. (2016). Integrating social and facial models of person perception: Converging and diverging dimensions. *Cognition*, *157*, 257–267. https://doi.org/10.1016/j.cognition.2016.09.006

Thompson, P., & Burr, D. (2009). Visual aftereffects. *Current Biology*, *19*(1), R11–R14.

Tigue, C. C., Borak, D. J., O'Connor, J. J. M., Schandl, C., & Feinberg, D. R. (2012). Voice pitch influences voting behavior. *Evolution and Human Behavior*, *33*(3), 210–216. https://doi.org/10.1016/j.evolhumbehav.2011.09.004

Titze, I. R. (1994). *Principles of voice production*. Englewood Cliffs, NJ: Prentice-Hall.

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social Attributions

from Faces: Determinants, Consequences, Accuracy, and Functional Significance. *Annual Review of Psychology*, *66*(1), 519–545. https://doi.org/10.1146/annurev-psych-113011-143831

Vukovic, J., Jones, B. C., Feinberg, D. R., DeBruine, L. M., Smith, F. G., Welling, L. L. M., & Little, A. C. (2011). Variation in perceptions of physical dominance and trustworthiness predicts individual differences in the effect of relationship context on women's preferences for masculine pitch in men's voices. British Journal of Psychology, 102, 37–48. http://dx.doi.org/10.1348/000712610X498750.

Wells, T., Baguley, T., Sergeant, M., & Dunn, A. (2013). Perceptions of Human Attractiveness Comprising Face and Voice Cues. *Archives of Sexual Behavior*, *42*(5), 805–811. https://doi.org/10.1007/s10508-012-0054-0

Winkielman, P., Halberstadt, J., Fazendeiro, T., & Catty, S. (2006). Prototypes Are Attractive Because They Are Easy on the Mind. *Psychological Science*, *17*(9), 799–806. https://doi.org/10.1111/j.1467-9280.2006.01785.x

Yovel, G., & Belin, P. (2013). A unified coding strategy for processing faces and voices. *Trends in Cognitive Sciences*, *17*(6), 263–271. https://doi.org/10.1016/j.tics.2013.04.004

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*(2, Pt.2), 1–27. https://doi.org/10.1037/h0025848

Zebrowitz, L. A., & Montepare, J. M. (2008). Social Psychological Face Perception: Why

Appearance Matters: Face Perception: Why Appearance Matters. *Social and Personality Psychology Compass*, *2*(3), 1497–1517. https://doi.org/10.1111/j.1751-9004.2008.00109.x

Zuckerman, M., Hodgins, H., & Miyake, K. (1990). The vocal attractiveness stereotype: Replication and elaboration. *Journal of Nonverbal Behavior*, *14*(2), 97–112. https://doi.org/10.1007/BF01670437

CHAPTER 2: ADAPTATION EFFECTS OF VOICE-PITCH ON ATTRACTIVENESS JUDGEMENTS

Ostrega, J., Little, A. C., & Feinberg, D. R. (2020). No Adaptation Effects of Voice-Pitch on Attractiveness Judgements.

PREFACE

After examining the results from the perceptual adaptation of voice pitch on normality judgments (Little et al., 2013), we were curious to assess if the same effects would be observed when we asked participants to evaluate attractiveness of speakers using the same manipulations and adaptation procedures. We expected normality and attractiveness judgments to produce similar effects, given our pre-existing understanding of the relationship between attractiveness and averageness (Bruckert et al., 2010; DeBruine et al., 2007).

Since previous studies had not considered women's perceptions of male voices we wanted to ensure that we studied a meaningful sample of women and men participants evaluating both male and female voices. This study sought to provide further evidence for perceptual adaptation of voice pitch and evaluate a new dimension in the perceptual adaptation literature using attractiveness judgements.

Our findings did not demonstrate that adaptation to voice pitch on attractiveness judgements produce an aftereffect. We did find a main effect of voice pitch on attractiveness judgements when women judged male voices. This result is likely a false

positive and does not eloquently represent perceptual effects of voice pitch manipulation on attractiveness judgements. These results encouraged a re-examination of perceptual adaptation results in voice research.

REFERENCES

Bruckert, L., Bestelmeyer, P., Latinus, M., Rouger, J., Charest, I., Rousselet, G. A., Kawahara, H., & Belin, P. (2010). Vocal Attractiveness Increases by Averaging. *Current Biology*, *20*(2), 116–120. https://doi.org/10.1016/j.cub.2009.11.034

DeBruine, L. M., Jones, B. C., Unger, L., Little, A. C., & Feinberg, D. R. (2007). Dissociating averageness and attractiveness: Attractive faces are not always average. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(6), 1420–1430. https://doi.org/10.1037/0096-1523.33.6.1420

Little, A. C., Feinberg, D. R., DeBruine, L. M., & Jones, B. C. (2013). Adaptation to Faces and Voices: Unimodal, Cross-Modal, and Sex-Specific Effects. *Psychological Science*, *24*(11), 2297–2305. https://doi.org/10.1177/0956797613493293

CHAPTER 2: ADAPTATION EFFECTS OF VOICE-PITCH ON ATTRACTIVENESS JUDGEMENTS

ABSTRACT

There are two competing theories about averageness and attractiveness. The average is attractive hypothesis states that people find average stimuli attractive because they closely resemble internal representations and are easy to process. The contrast hypothesis states that attractiveness depends on contrast from average such that exaggerated traits of stimuli in one direction increases attractiveness and exaggerated traits of stimuli in the opposite direction decrease attractiveness. Studies on facial attractiveness show that after adaptation to unattractive faces, unattractive faces become more normal looking, and more attractive, whereas attractive faces become less normal and less attractive. Voices are thought to be processed in a similar fashion to faces, and the strength of preferences for sex-typical voice and face features are positively correlated. After adaptation to female voices with sex-typical or sex-atypical voice pitch, the voices that participants adapted to sounded more normal. We tested if adaptation to sex-typical and sex-atypical voice pitch (the primary correlate of voice attractiveness) influences the attractiveness of voices. Using identical stimuli and paradigms as prior work showing adaptation effects of voice pitch on normality judgments, we tested for adaptation effects of high and low-pitched voices for both male and female vocalizers. We observed one potential adaptation effect when women listened to men's voices, consistent with the "averageness is attractive" hypothesis, but there were no effects in any other condition. Overall,

adaptation based on sex-typicality did not change attractiveness judgments consistently and the one effect seen here may reflect a false-positive result.

INTRODUCTION

There are two competing hypotheses as to why average voices and faces are thought to be more attractive than average. One theory is that the more average something is, the more attractive it is because it is closer to our internal prototype (for review, see Belin, Bestelmeyer, Latinus, & Watson, 2011), which subsequently increases perceptual fluency or ease of processing of such voices (Belin et al., 2011; Bruckert et al., 2010; Latinus & Belin, 2011). The other theory expands on the averageness is attractive hypothesis suggesting that moving away from averageness along particular dimensions can increase attractiveness even more (DeBruine et al., 2007).

Many studies have tested if increased exposure to a stimulus increases perceptions of its averageness. Increased exposure to a stimulus has also been examined to influence preferences for the stimulus. Previous exposure to objects, music, faces and voices has been shown to increase positive affect towards those stimuli (Peretz et al., 1998; Rhodes et al., 2001; Zajonc, 1968). Repeated exposure to faces and specific face characteristics increases liking and attractiveness judgements of those faces (Webster et al., 2004). After repeated exposure participants like unfamiliar auditory stimuli, such as melodies, more than without repeated exposure (McDermott, 2012; Peretz et al., 1998). Celebrity voices that are recognized by participants are perceived as more pleasant sounding than unrecognized celebrity voices (Vinney & Vinney, 2017). Additionally, participants rate their own voices as sounding more attractive in comparison to other voices and rate their own voice as more attractive compared to how others rate their voice (Hughes & Harrison, 2013; Peng et al., 2019). It is possible that increased exposure to voices and

other auditory stimuli may lead to mere exposure effects. Auditory stimuli previously heard may be liked more or sound more attractive because they reflect recent updates to our internal prototypes. Listeners have expectations about how male and female voices sound (Babel et al., 2014; Babel & McGuire, 2015; Casado & Brunellière, 2016; Little et al., 2013) and the attractiveness of both male and female voices correlates with how close voices are to the average (Babel & McGuire, 2015; Bruckert et al., 2010). When voice morphs were created using many voices to represent the average, they were judged as sounding more attractive and the more voices used in the average, the more attractive they sounded (Bruckert et al., 2010).

Facial attractiveness judgments may depend on the direction from which they diverge from average, and not just the distance from average (DeBruine et al., 2007). Adaptation to attractive faces makes them look more normal (DeBruine et al 2007; Little et al., 2013) and less attractive (DeBruine et al 2007). Whereas unattractive faces become less normal (DeBruine et al 2007; Little et al., 2013), but they do not change in attractiveness (DeBruine et al 2007). Adaptation to unattractive faces makes unattractive faces seem more normal and more attractive, attractive faces become less normal and more attractive (DeBruine et al 2007). There is some evidence for the aforementioned findings that the direction from averageness influences attractiveness in voices as well as faces. For example, people tend to prefer lower than average male voice pitch and higher than average female voice pitch (Apicella & Feinberg, 2009; Collins, 2000; Feinberg et al., 2005; Jones et al., 2008, 2010; Puts et al., 2006; Re et al., 2012; Vukovic et al., 2008; Zhang et al., 2018).

Perceptual aftereffects for voices have been observed for individual identity (Belin & Zatorre, 2003), familiarity (Zäske et al., 2010), and gender (Zäske et al., 2009). When people were adapted to female voices, they rated subsequent stimuli as sounding more male and when people were adapted to male voices, they rated subsequent stimuli as sounding more female. This effect was more pronounced when test stimuli were ambiguous along the sex-typicality continuum (Zäske et al., 2009); voices that possessed both feminine and masculine voice attributes, rather than mostly feminine or masculine attributes (Zäske et al., 2009).

Zäske et al., (2009) found that adapting to male and female voices changed subsequent perceptions of maleness and femaleness. Voice pitch is the primary acoustic differentiator of male and female voices (Bachorowski & Owren, 1999). Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. As such, other work has found that after exposure to high pitched female voices, participants later judged high pitched female voices as sounding more normal (Little et al., 2013). If, however, participants were exposed to low pitched female voices, they later judged low pitched female voices as sounding more normal. The same aftereffect was not present when participants were adapted to either high- or low-pitched female voice and then asked to judge the normality of male voices. In this study, male voices were never used as an adaptive stimulus to test for aftereffects (Little et al., 2013) and a focus on one sex as adapting stimuli is an issue in other work on face aftereffects (DeBruine et al., 2007).

While averageness is preferred in both voices and faces, Zäske et al. (2020) found that voice and face averageness are not correlated. Nevertheless, adaptation to voice pitch alters subsequent face perception (Little et al., 2013). This suggests that it is unlikely that preferences for averageness are based on some intrinsic qualities of the individual being perceived. Cognitive biases in voice and face perception such as preferences for perceptually fluent stimuli are more likely to explain these findings.

From an evolutionary perspective it is possible that selection pressures for a parenting partner might have influenced how men and women have adapted to voice signals in selecting high-quality parenting partners (Trivers, 1972). Participants might be more sensitive to the voice pitch of the sex of voice that belong to the sex of their preferred parenting partner compared to other voices. Popular evolutionary theories also highlight male dominance as the driver of sex differences in voice pitch (Aung & Puts, 2020). Consequently, men might be more sensitive to the changes in the voice pitch of competing males (Saxton et al., 2016).

Predictions for Averageness is Attractive Hypothesis

The averageness is attractive hypothesis suggests that after exposure to high or low pitch, preference for stimuli at the same pitch level as the adapting stimuli will increase. For women's voices when people are exposed to high pitch (making it more normal), they will prefer high pitch more than before exposure. When they hear low pitch more (making it more normal), they will prefer high pitch less than before exposure. For men's voices, when people are exposed to low pitch (making it more normal), they will

prefer low pitch more than before exposure. When they hear high pitch more (making it more normal), they will prefer low pitch less than before exposure.

Predictions for the Contrast Hypothesis

Following the contrast hypothesis (DeBruine et al., 2007), which suggests the *direction* from average qualifies the averageness is attractive hypothesis, we can propose two predictions. If voices are already attractive, and we are exposed to them more, their attractiveness diminishes, they become less attractive sounding over time. On the other hand, if the voices start out less attractive, the more we hear them, the more attractive they become. Since many studies have shown that people prefer low pitch in men's voices, and high pitch in women's voices (Apicella & Feinberg, 2009; Collins, 2000; Feinberg et al., 2005; Jones et al., 2008, 2010; Puts et al., 2006; Re et al., 2012; Vukovic et al., 2008; Zhang et al., 2018), we predict exposure to low pitch in men's voices would make low pitch less attractive and high pitch more attractive. Conversely, exposure to high pitch in men's voices would make high pitch more attractive, and low pitch less attractive. If people prefer high pitch in women's voices, exposure to high pitch would make high pitch voices sound less attractive and low-pitched voices more attractive. Exposure to low pitched women's voices would make lower pitched voices sound more attractive. Prior work on the relationship between averageness and attractiveness have not found any sex differences in preferences for averageness (Bruckert et al., 2010).

METHODS

Stimuli

We used stimuli from a sample of undergraduate student voices Little et al. (2013). Voice clips of individuals speaking common English vowel sounds (International Phonetic Alphabet: ɑ/, /ɛ/, /i/, /o/, /u/) that had the closest mean pitch values to the groups average for each sex in a larger sample of about 100 women's voices and 50 men's voices. This number of stimuli has been shown to successfully change perceptions of voice attractiveness (Borkowska & Pawlowski 2011; Feinberg et al 2005; 2006; 2008).

Experiment Phase	Sex of Voice	Pitch Manipulation 20 Hz	Mean (Hz)	Standard deviation (Hz)	Min (Hz)	Max (Hz)
Adaptation Phase	Female	Lowered	182.61	2.07	181.14	184.07
		Raised	232.62	1.96	231.24	234.01
	Male	Lowered	90.3	0.15	90.19	90.4
		Raised	130.11	0.04	130.09	130.14
Test Phase	Female	Lowered	181.51	7.22	168.63	190.49
		Raised	231.53	6.87	219.8	239.94
	Male	Lowered	90.28	0.32	89.95	90.84
		Raised	130.12	0.22	129.82	130.55

TABLE 1 shows that all the voices in both the adaptation and test conditions that were lowered were lower than all the voices that were raised for both sexes of vocalizer. The means of the voices for each manipulation for each sex of voice across test phases are nearly identical.

We manipulated voices by increasing or decreasing the voice pitch of 10 male and 10 female voices by 20 Hz using Praat's pitch-synchronous overlap add (PSOLA) algorithm (Boersma & Weenink, 2010) (see Feinberg et al., 2005), leaving duration and formant frequencies unmanipulated. This manipulation has successfully been used in other studies that examine human voice attractiveness (Feinberg et al., 2006, 2008; Jones et al., 2008, 2010; Puts et al., 2006; Vukovic et al., 2008) along with studies of mate

preference, dominance, and voice quality of other mammalian species (Ghazanfar et al., 2007; Reby et al., 2005). Altogether there were 4 stimuli sets, containing 5 pairs of stimuli per set and 2 stimuli sets per sex. Each stimulus set contained 5 pairs of the same voice (high and low pitch manipulation), the average stimuli duration was 6 seconds. The total exposure time was 60 seconds and achieved by looping the series of sounds until 60 seconds terminated. Each stimulus set contained a distinct set of 5 voices. There were 2 sets of 5 female voices and 2 sets of 5 male voices. We acknowledge that this is a small stimuli set. Considering Little et al. (2013) and previous studies (DeBruine et al., 2007; Zäske et al., 2009, 2010) have used this sample size and found significant adaptation effects, we feel justified that this is a reasonable sample size to observe any significant effects.

Participants

Protocols for this study were approved by the McMaster University Research Ethics Board. We aimed to double the sample size from the experiment by Little et al. (2013). Prior to analyses, we excluded all incomplete surveys. Participants were recruited from the McMaster University on-line subject pool (SONA) of undergraduate students and were compensated with extra course credit or cash payment ($10 per hour pro rata) for participation.

Participants identified their sex and gender by answering a 3-part questionnaire that asked participants their sex at birth, the gender they currently live as, and the gender they identified as on the day of testing (Fraser, 2018). One hundred fifty-three participants identified the gender they currently live as women, 84 as men, 2 as nonbinary or 2-

spirited, and 2 skipped the question. Participant age (17-26) was self-reported. Sexual orientation was also self-reported using a modified version of the Kinsey scale (Kinsey et al., 1953) that included options for “asexual”, and “other”. Whilst these data were recorded, and analyzed, we did not have any specific hypotheses that these groups would differ in their responses.

Procedure

Responder Gender	Voice Stimuli							
	Female				Male			
	Same		Novel		Same		Novel	
	High	Low	High	Low	High	Low	High	Low
Woman	19	8	13	15	26	25	22	25
Man	24	7	16	16	5	4	7	5
Non-binary	1	0	0	0	0	1	0	0

TABLE 2 shows the number of participants who responded across all 8 conditions. With each person responding to female or male voice stimuli, adapting to high- or low- pitched voices, and to the same or novel voices to those they were asked to judge in the pre- and post- adaptation responses.

Participants were assigned one of the 8 adaptation conditions in a counterbalanced order and were adapted to either high-pitched or low-pitched manipulation of the same identity of either male or female speakers. In other words, participants only complete one version of the experiment, and versions varied between participants.

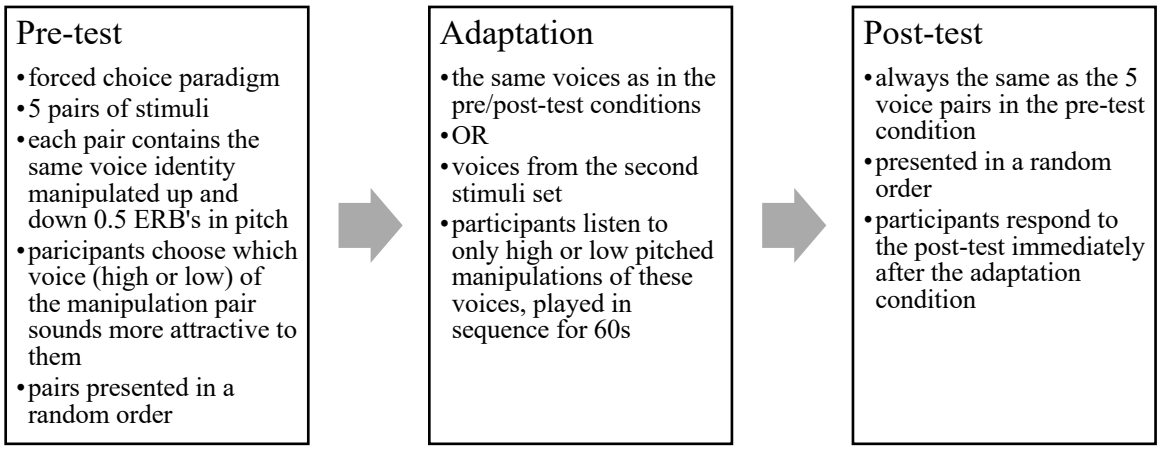

FIGURE 1 displays the three phases of our experiment including the pre-test, adaptation, and post-test.

There were three phases of the experiment: pre-test, adaptation, and post-test. In the pretest, we assessed participants' preferences for high/low-pitched voices by choosing among 5 pairs of raised and lowered pitch voices for which was more attractive using a forced choice paradigm. Participants were presented with pairs of the same voice raised and lowered in pitch. Pairs of voices were presented in random orders. In half of the conditions, the same voice pair stimuli were used during the adaptation phase and in the other half participants were adapted to novel stimuli from the other stimuli set to test if any adaptation effects generalized to the stimulus category rather than specific to voice identity. We dropped this later variable after finding it had no effect on any analysis.

In the adaptation phase, participants listened to a series of either increased-pitch or decreased-pitch voices of same or novel stimuli, looped for 1 minute. Following the adaptation phase, participants immediately repeated the pre-test procedure. The post-test phase was followed by a brief questionnaire regarding the participants self-report of sexual-orientation, sex, gender, age, and whether English was their first language.

Data Processing

We compiled our data using R version 4.1.2 (R Core Team, 2021), we organized our data using tidyverse 1.3.1 (Wickham et al., 2019) and used multilevel linear regression with lmerTest 3.1.3 (Kuznetsova et al., 2017) and jtools 2.1.4 (Long, 2020) to analyze our data. For each model, random intercepts were specified for each audio stimulus and for each participant to control for nonindependence of responses to the same stimulus and from the same participant, and random slopes were entered for each within-subject or within-stimulus variable at each corresponding intercept, as suggested in Barr et al. (2013) and Barr (2013).

Our response variable was preference for sex typicality. Here, sex typical means they preferred lower pitch among men's voices and higher pitch among women's voices. Our fixed effects variables were an interaction between whether people adapted to sex typical and sex atypical voices, their preference before and after adaptation.

Pre- and post-test responses to voice pairs of sex typical (high pitched for female and low pitched for male) and sex atypical (low pitched for female and high pitched for male) voice manipulations were compared with the exposure of participants to sex typical and sex atypical voices. The variable regarding whether the same voices were in the adaptation (same or novel voices) and test phases yielded no significant effects or interactions, so it was excluded from these analyses to simplify interpretation. All data and analysis code are publicly available at https://osf.io/93tkx/.

RESULTS

Separate models were run for women responding to female voice stimuli, women responding to male voice stimuli, men responding to female voice stimuli and men responding to male voice stimuli. The R code for the full models is in the supplementary material.

Fixed effects from the regression analyses are summarized in Table 3. (Women responding to female voice stimuli), Table 4. (Women responding to male voice stimuli), Table 5. (Men responding to female voice stimuli), Table 6. (Men responding to male voice stimuli), and Figure 2. Adaptation effects were only found when women responded to male voice stimuli and were not affected by the sex-typicality of the adapting stimuli.

	Odds Ratio	**Lower CI**	**Upper CI**	**z**	**p**
Intercept	1.695	0.968	2.421	4.572	<0.0001***
PrePost	0.226	-0.836	1.288	0.417	0.677
Adaptation to Sex Typical	-0.925	-1.767	-0.084	-2.155	0.031*
PrePost x Adaptation to Sex Typical	-0.020	-1.220	1.181	-0.032	0.975

TABLE 3 Results of our analysis for women responding to female voice stimuli. Women preferred female voices with higher pitch to those in lower pitch. All ratings were slightly lower when people heard lower pitched voices, but this did not affect any change in preferences over time.
*p<0.05; **p<=0.01; ***p<=0.0001

	Odds Ratio	**Lower CI**	**Upper CI**	**z**	**p**
Intercept	0.061	-0.316	0.439	0.319	0.750
PrePost	-0.539	-0.972	-0.107	-2.444	0.015*
Adaptation to Sex Typical	0.296	-0.234	0.825	1.093	0.274
PrePost x Adaptation to Sex Typical	1.002	0.419	1.585	3.369	0.001**

TABLE 4 Results of our analysis for women responding to male voice stimuli. After exposure to low pitched voices, women preferred high pitched men's voices more. After exposure to high pitched voices, women preferred low pitch more.
*p<0.05; **p<=0.01; ***p<=0.0001

	Odds Ratio	**Lower CI**	**Upper CI**	**z**	**p**
Intercept	1.115	0.633	1.596	4.540	<0.001***
PrePost	0.016	-0.683	0.715	0.046	0.964
Adaptation to Sex Typical	0.525	-0.059	1.108	1.762	0.078
PrePost x Adaptation to Sex Typical	-0.260	-1.063	0.543	-0.634	0.526

TABLE 5 Results of our analysis for men responding to female voice stimuli. After exposure to high pitch, men did not change their preferences for high pitched voices. After exposure to low pitch, men preferred women with lower pitch slightly more, but this was not significant.
*p<0.05; **p<=0.01; ***p<=0.0001

	Odds Ratio	**Lower CI**	**Upper CI**	**z**	**p**
Intercept	0.228	-0.428	0.884	0.682	0.495
PrePost	0.403	-0.501	1.307	0.874	0.382
Adaptation to Sex Typical	-0.811	-1.867	0.244	-1.507	0.132
PrePost x Adaptation to Sex Typical	0.438	-0.824	1.699	0.680	0.496

TABLE 6 Results of our analysis for men responding to male voice stimuli. After hearing both low pitched and high-pitched voices, men preferred lower pitch in men's voices more.
*p<0.05; **p<=0.01; ***p<=0.0001

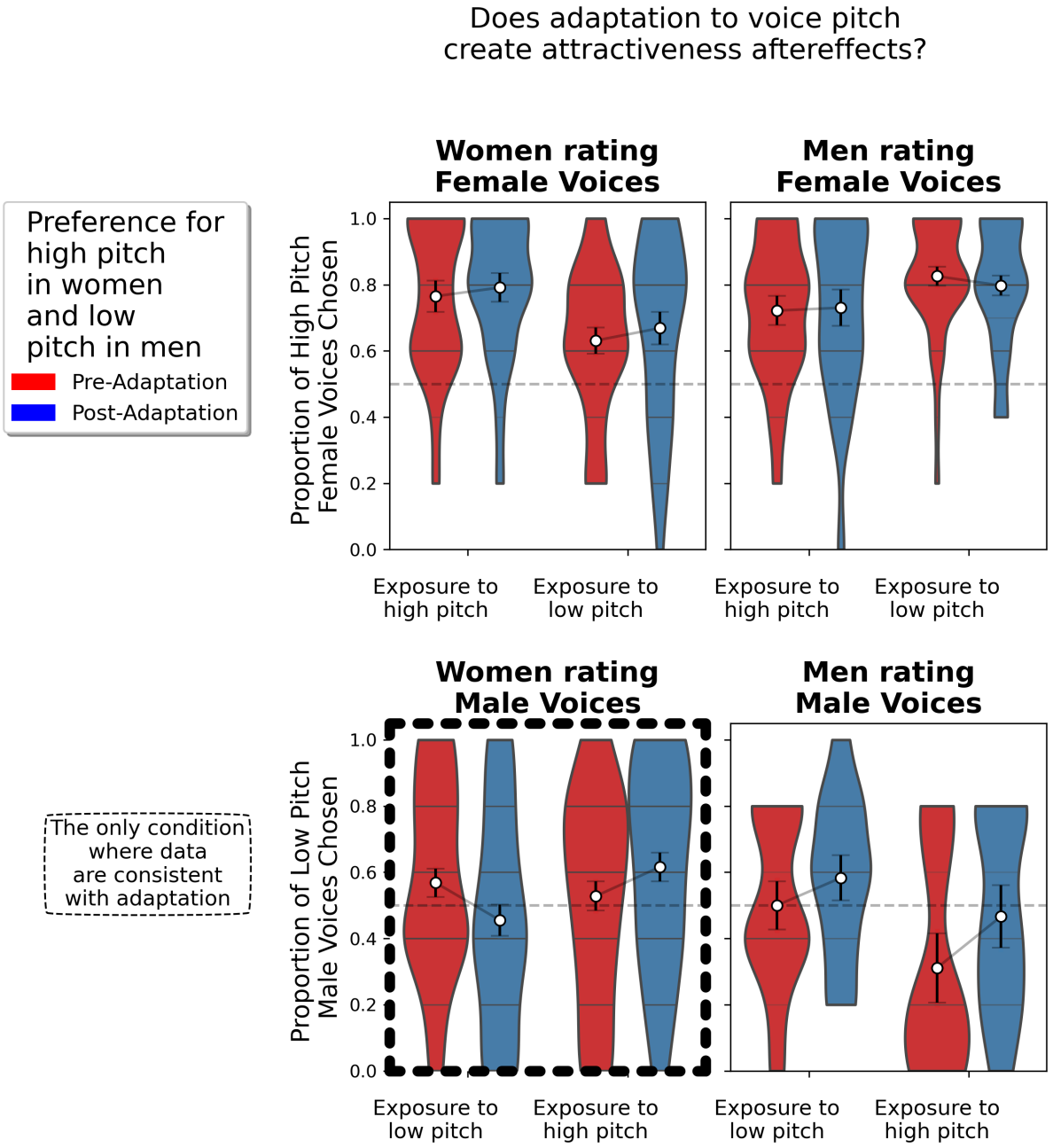

FIGURE 2 displays violin plots. The white dot in the centre of each violin represents the mean, and the bars are standard error bars. This plot displays the proportion high pitched female voices chosen (top panel), and the proportion of low-pitched men's voices chosen (bottom panel). The left panels show data from people who identify as women, the right panels display data from people who identify as men.

DISCUSSION

We tested if adaptation to high- or low-pitched voices influenced attractiveness judgements of both female and male speakers. We expected to find results consistent with those in perceptual adaptation to voice pitch on normality judgements (Little et al., 2013) and the averageness is attractive hypothesis (Bruckert et al 2010). One possible adaptation effect was observed when women responded to male voices, but no other effects.

We found data consistent with the contrast hypothesis (Debruine et al., 2007) when women were responding to men's voices. When women were adapted to low-pitched male voices, they preferred low-pitched male voices less and high-pitched male voices more. When women adapted to high-pitched male voices, they preferred high-pitched male voices less and low-pitched male voices more. Here, the adapting stimuli made the corresponding test voices sound less attractive, suggesting that the direction from average matters, and distance from average is not the only thing that predicts voice attractiveness. However, we did not find any evidence for the contrast hypothesis in any other condition.

When men rated women's voices, and were exposed to high pitch, no adaptation effects were found, which is consistent with (Debruine et al., 2007), who found adapting to attractive faces had little effect on subsequent ratings. However, when men were exposed to low pitch in women's voices, which was less attractive in this and other studies (Apicella & Feinberg, 2009; Jones et al., 2008, 2010; Re et al., 2012; Vukovic et al., 2008; Zhang et al., 2018), they increased their preferences for low pitch very slightly, but not enough to reach statistical significance, rendering this a null result. Had this been

significant, it would be consistent with the averageness is attractive hypothesis (Bruckert et al 2010).

Our results are surprising since we used the same stimuli and procedure as in Little et al. (2013), who demonstrated that adaptation to high and low female voice pitch produced normality aftereffects in female voices. Importantly Little et al. (2013) did not test male voices as adaptive stimuli. Attractiveness aftereffects have been found in the past for female faces only (DeBruine et al., 2007), and only when people adapted to unattractive faces. Accordingly, even the evidence that adapting to faces changes attractiveness ratings is not internally consistent in prior work.

Both voice and face perception are thought to share the same underlying coding strategy (Yovel & Belin, 2013), therefore, we would have anticipated similar aftereffects in voices. Our results for women responding to male voices were in a direction consistent with the contrast hypothesis, which could be consistent with evolutionary theories where female choice has an impact on sexual selection in voices (see Feinberg et al 2018; 2019), but not consistent with potentially gender-biased theories where only male dominance matters for evolution, and female choice is not important (Puts & Aung, 2019).

It is unlikely that our null results were due to any design or execution error on our part. We followed the exact study design as used in Little et al. (2013), using the same voice stimuli and the same timing used for the pre-test, adaptation, and post-test phases. Our sample of responders was also more than double that of Little et al. (2013) making it unlikely that our results reflect insufficient power and are false negatives. Furthermore, we did find an overall preference for high-pitched female and low-pitched male voices in

all pre and post responses regardless of the adaptation condition, gender of responder or sexual orientation of responder. This is consistent with previous findings in voice research on voice pitch and attractiveness judgments (Apicella & Feinberg, 2009; Collins, 2000; Feinberg et al., 2008; Jones et al., 2008, 2010; Puts et al., 2006; Re et al., 2012; Vukovic et al., 2008; Zhang et al., 2018, 2019). It is also consistent with work on homosexual men's perception of voice attractiveness of both men and women's voices (Valentová et al., 2013; Zhang et al., 2018, 2019). Nevertheless, it is also possible that uneven sample sizes also affected our data, but were this the case, we would have expected adaptation effects on attractiveness in the 3/4 conditions with larger samples than Little et al (2013). This did not happen.

The finding that adaptation only occurred when women were rating men's voices may be a false positive. However, there is also a plausible explanation from an evolutionary perspective. Selective pressures to invest in a high quality parenting partner (Trivers, 1972) might have resulted in adaptation. Therefore, women may be more sensitive to detecting changes in men's voices. Popular theories of the evolution of sexually dimorphic voice pitch downplay the role of women's choices, and emphasize male dominance as the driver of evolution of sex differences in voice pitch (Aung & Puts, 2020). If women's choices are not important, as Aung and Puts suggest, then we would not expect to find adaptation effects only when women respond to men's voices. One study has predicted that men may be more sensitive to the changes in the fundamental frequency of competing males (Saxton et al., 2016). If we are evoking an evolutionary explanation here, it is unclear why women's choices are so critical. Alternatively, the

observed adaptation effect is a false positive, considering we found no adaptation effects in 3/4 of the conditions.

A possible explanation for the lack of main effects for responding to female voices could be ceiling effects of attractiveness judgments. Since high-pitched female voices were already strongly preferred over low-pitched female voices using our stimuli pairs, it might have been difficult to detect adaptation effects to female voices in our sample and our results may be indicative of a false negative. Ceiling effects were not observed in any of our conditions (see Figure 1).

Another potential explanation of our null results could be that we studied a single feature, rather than holistic manipulations. However, this single feature adaption did alter normality judgements in Little et al. (2013). Other work has shown single features in the visual domain, such as eye spacing in faces, are sufficient to produce visual aftereffects (Little et al., 2005). Therefore, while it is a possible explanation of our results, we are uncertain why holistic manipulations would be required to yield aftereffects in voice attractiveness, but not in voice normality judgements using the same stimuli set.

It is important to note that the effect for women adapting to male voices although significant is weak. Other perceptual adaptation effects that are observed, particularly visual aftereffects occur over a short time scale, and affect most participants (Anstis et al., 1998; Thompson & Burr, 2009; Webster et al., 2004). They are not difficult to detect, nor do they require large numbers of stimuli (Anstis et al., 1998; Thompson & Burr, 2009; Webster et al., 2004). It may be that auditory perceptual adaptations, particularly to voices may be more difficult to detect than visual adaptations. Although adaptation to some

voice characteristics may result in more dramatic perceptual adaptation effects than others, perceptual adaptation to all voice stimuli should be evaluated and examined critically, and with reference to potential file-drawer issues. In summary, we found little evidence for vocal attractiveness aftereffects. We call for a re-examination of previous perceptual aftereffects and research on averageness and voice attractiveness.

REFERENCES

Anstis, S., Verstraten, F. A., & Mather, G. (1998). The motion aftereffect. *Trends in Cognitive Sciences*, *2*(3), 111–117.

Apicella, C. L., & Feinberg, D. R. (2009). Voice pitch alters mate-choice-relevant perception in hunter–gatherers. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1659), 1077–1082. https://doi.org/10.1098/rspb.2008.1542

Aung, T., & Puts, D. (2020). Voice pitch: A window into the communication of social power. *Current Opinion in Psychology*, *33*, 154–161. https://doi.org/10.1016/j.copsyc.2019.07.028

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Babel, M., & McGuire, G. (2015). Perceptual Fluency and Judgments of Vocal Aesthetics and Stereotypicality. *Cognitive Science*, *39*(4), 766–787. https://doi.org/10.1111/cogs.12179

Babel, M., McGuire, G., & King, J. (2014). Towards a More Nuanced View of Vocal Attractiveness. *PLoS ONE*, *9*(2), e88616. https://doi.org/10.1371/journal.pone.0088616

Bachorowski, J.-A., & Owren, M. J. (1999). Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. *The Journal of the Acoustical Society of America*, *106*(2), 1054–1063. https://doi.org/10.1121/1.427115

Belin, P., Bestelmeyer, P. E. G., Latinus, M., & Watson, R. (2011). Understanding Voice Perception: Understanding voice perception. *British Journal of Psychology*,*102*(4), 711–725. https://doi.org/10.1111/j.2044-8295.2011.02041.x

Belin, P., & Zatorre, R. J. (2003). Adaptation to speaker's voice in right anterior temporal lobe: *NeuroReport*, *14*(16), 2105–2109. https://doi.org/10.1097/00001756-200311140-00019

Boersma, P., & Weenink, D. (2010). *Praat: Doing phonetics by computer*. www.praat.org

Bruckert, L., Bestelmeyer, P., Latinus, M., Rouger, J., Charest, I., Rousselet, G. A., Kawahara, H., & Belin, P. (2010). Vocal Attractiveness Increases by Averaging. *Current Biology*, *20*(2), 116–120. https://doi.org/10.1016/j.cub.2009.11.034

Casado, A., & Brunellière, A. (2016). The influence of sex information into spoken words: A mismatch negativity (MMN) study. *Brain Research*, *1650*, 73–83. https://doi.org/10.1016/j.brainres.2016.08.039

Collins, S. A. (2000). Men's voices and women's choices. *Animal Behaviour*, *60*(6), 773–780. https://doi.org/10.1006/anbe.2000.1523

DeBruine, L. M., Jones, B. C., Little, A. C., Boothroyd, L. G., Perrett, D. I., Penton-Voak, I. S., Cooper, P. A., Penke, L., Feinberg, D. R., & Tiddeman, B. P. (2006). Correlated preferences for facial masculinity and ideal or actual partner's masculinity. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1592), 1355–1360. https://doi.org/10.1098/rspb.2005.3445

DeBruine, L. M., Jones, B. C., Unger, L., Little, A. C., & Feinberg, D. R. (2007).

Dissociating averageness and attractiveness: Attractive faces are not always average. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(6), 1420–1430. https://doi.org/10.1037/0096-1523.33.6.1420

Dekam, C.-R. J., Morey, A., & Tuma, L. (2008). *The Effects of Voice Pitch on Perceptions of Attractiveness: Do You Sound Hot or Not?* 6.

Feinberg, D. R., DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2008). The Role of Femininity and Averageness of Voice Pitch in Aesthetic Judgments of Women's Voices. *Perception*, *37*(4), 615–623. https://doi.org/10.1068/p5514

Feinberg, D. R., Jones, B. C., Law Smith, M. J., Moore, F. R., DeBruine, L. M., Cornwell, R. E., Hillier, S. G., & Perrett, D. I. (2006). Menstrual cycle, trait estrogen level, and masculinity preferences in the human voice. *Hormones and Behavior*, *49*(2), 215–222. https://doi.org/10.1016/j.yhbeh.2005.07.004

Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour*, *69*(3), 561–568. https://doi.org/10.1016/j.anbehav.2004.06.012

Fraser, G. (2018). Evaluating inclusive gender identity measures for use in quantitative psychological research. *Psychology & Sexuality*, *9*(4), 343–357. https://doi.org/10.1080/19419899.2018.1497693

Ghazanfar, A. A., Turesson, H. K., Maier, J. X., van Dinther, R., Patterson, R. D., & Logothetis, N. K. (2007). Vocal-Tract Resonances as Indexical Cues in Rhesus

Monkeys. *Current Biology*, *17*(5), 425–430. https://doi.org/10.1016/j.cub.2007.01.029

Hughes, S. M., & Harrison, M. A. (2013). I like My Voice Better: Self-Enhancement Bias in Perceptions of Voice Attractiveness. *Perception*, *42*(9), 941–949. https://doi.org/10.1068/p7526

Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C., & Vukovic, J. (2008). Integrating cues of social interest and voice pitch in men's preferences for women's voices. *Biology Letters*, *4*(2), 192–194. https://doi.org/10.1098/rsbl.2007.0626

Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C., & Vukovic, J. (2010). A domain-specific opposite-sex bias in human preferences for manipulated voice pitch. *Animal Behaviour*, *79*(1), 57–62. https://doi.org/10.1016/j.anbehav.2009.10.003

Kinsey, A. C., Pomeroy, W. B., Martin, C. E., & Gebhard, P. H. (1953). *Sexual behavior in the human female*. Indiana University Press.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Latinus, M., & Belin, P. (2011). Anti-Voice Adaptation Suggests Prototype-Based Coding of Voice Identity. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00175

Little, A. C., DeBruine, L. M., & Jones, B. C. (2005). Sex-contingent face after-effects

suggest distinct neural populations code male and female faces. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1578), 2283–2287. https://doi.org/10.1098/rspb.2005.3220

Little, A. C., Feinberg, D. R., DeBruine, L. M., & Jones, B. C. (2013). Adaptation to Faces and Voices: Unimodal, Cross-Modal, and Sex-Specific Effects. *Psychological Science*, *24*(11), 2297–2305. https://doi.org/10.1177/0956797613493293

Long, J. A. (2020). *jtools: Analysis and Presentation of Social Scientific Data*. https://cran.r-project.org/package=jtools

McDermott, J. H. (2012). Auditory Preferences and Aesthetics. In *Neuroscience of Preference and Choice* (pp. 227–256). Elsevier. https://doi.org/10.1016/B978-0-12-381431-9.00020-6

Peng, Z., Wang, Y., Meng, L., Liu, H., & Hu, Z. (2019). One's own and similar voices are more attractive than other voices. *Australian Journal of Psychology*, *71*(3), 212–222. https://doi.org/10.1111/ajpy.12235

Peretz, I., Gaudreau, D., & Bonnel, A.-M. (1998). Exposure effects on music preference and recognition. *Memory & Cognition*, *26*(5), 884–902. https://doi.org/10.3758/BF03201171

Pisanski, K., & Bryant, G. A. (2019). The Evolution of Voice Perception. *The Oxford handbook of voice studies*, 269-300.

Pisanski, K., & Feinberg, D. R. (2018). *27.1 The Interdisciplinary Nature of Voice Attractiveness Research*. 21.

Puts, D. A., Gaulin, S. J. C., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior*, *27*(4), 283–296. https://doi.org/10.1016/j.evolhumbehav.2005.11.003

R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Re, D. E., O'Connor, J. J. M., Bennett, P. J., & Feinberg, D. R. (2012). Preferences for Very Low and Very High Voice Pitch in Humans. *PLoS ONE*, *7*(3), e32719. https://doi.org/10.1371/journal.pone.0032719

Reby, D., McComb, K., Cargnelutti, B., Darwin, C., Fitch, W. T., & Clutton-Brock, T. (2005). Red deer stags use formants as assessment cues during intrasexual agonistic interactions. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1566), 941–947. https://doi.org/10.1098/rspb.2004.2954

Rhodes, G., Halberstadt, J., & Brajkovich, G. (2001). Generalization of Mere Exposure Effects to Averaged Composite Faces. *Social Cognition*, *19*(1), 57–70. https://doi.org/10.1521/soco.19.1.57.18961

Saxton, T. K., Mackey, L. L., McCarty, K., & Neave, N. (2016). A lover or a fighter? Opposing sexual selection pressures on men's vocal pitch and facial hair. *Behavioral Ecology*, *27*(2), 512–519. https://doi.org/10.1093/beheco/arv178

Shirazi, T. N., Puts, D. A., & Escasa-Dorne, M. J. (2018). Filipino Women's Preferences for Male Voice Pitch: Intra-Individual, Life History, and Hormonal Predictors. *Adaptive Human Behavior and Physiology*, *4*(2), 188–206. https://doi.org/10.1007/s40750-018-0087-2

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, *22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Simmons, L. W., Peters, M., & Rhodes, G. (2011). Low Pitched Voices Are Perceived as Masculine and Attractive but Do They Predict Semen Quality in Men? *PLoS ONE*, *6*(12), e29271. https://doi.org/10.1371/journal.pone.0029271

Thompson, P., & Burr, D. (2009). Visual aftereffects. *Current Biology*, *19*(1), R11–R14.

Trivers, R. L. (1972). Parental investment and sexual selection. In Sexual selection and the descent of man. *Routledge*, 136–179.

Valentová, J., Roberts, S. C., & Havlíček, J. (2013). Preferences for Facial and Vocal Masculinity in Homosexual Men: The Role of Relationship Status, Sexual Restrictiveness, and Self-Perceived Masculinity. *Perception*, *42*(2), 187–197. https://doi.org/10.1068/p6909

Valentová, J. V., Varella, M. A. C., Havlíček, J., & Kleisner, K. (2017). Positive association between vocal and facial attractiveness in women but not in men: A cross-cultural study. *Behavioural Processes*, *135*, 95–100. https://doi.org/10.1016/j.beproc.2016.12.005

Vinney, C., & Vinney, L. A. (2017). That Sounds Familiar: The Relationship Between Listeners' Recognition of Celebrity Voices, Perceptions of Vocal Pleasantness, and Engagement with Media. *Journal of Radio & Audio Media*, *24*(2), 320–338. https://doi.org/10.1080/19376529.2017.1346659

Vukovic, J., Feinberg, D. R., Jones, B. C., DeBruine, L. M., Welling, L. L. M., Little, A. C., & Smith, F. G. (2008). Self-rated attractiveness predicts individual differences in women's preferences for masculine men's voices. *Personality and Individual Differences*, *45*(6), 451–456. https://doi.org/10.1016/j.paid.2008.05.013

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature*, *428*(6982), 557–561.https://doi.org/10.1038/nature02420

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., & Grolemund, G. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Winkielman, P., Halberstadt, J., Fazendeiro, T., & Catty, S. (2006). Prototypes Are Attractive Because They Are Easy on the Mind. *Psychological Science*, *17*(9), 799–806. https://doi.org/10.1111/j.1467-9280.2006.01785.x

Yovel, G., & Belin, P. (2013). A unified coding strategy for processing faces and voices. *Trends in Cognitive Sciences*, *17*(6), 263–271. https://doi.org/10.1016/j.tics.2013.04.004

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*(2, Pt.2), 1–27. https://doi.org/10.1037/h0025848

Zäske, R., Schweinberger, S. R., Kaufmann, J. M., & Kawahara, H. (2009). In the ear of the beholder: Neural correlates of adaptation to voice gender. *European Journal of Neuroscience*, *30*(3), 527–534. https://doi.org/10.1111/j.1460-9568.2009.06839.x

Zäske, R., Schweinberger, S. R., & Kawahara, H. (2010). Voice aftereffects of adaptation to speaker identity. *Hearing Research*, *268*(1–2), 38–45.

https://doi.org/10.1016/j.heares.2010.04.011

Zäske, R., Skuk, V. G., & Schweinberger, S. R. (2020). Attractiveness and distinctiveness between speakers' voices in naturalistic speech and their faces are uncorrelated. *Royal Society Open Science*, *7*(12), 201244. https://doi.org/10.1098/rsos.201244

Zhang, J., Zheng, L., & Zheng, Y. (2018). Consistency in preferences for masculinity in faces, bodies, voices, and personality characteristics among homosexual men in China. *Personality and Individual Differences*, *134*, 137–142. https://doi.org/10.1016/j.paid.2018.06.009

Zhang, J., Zheng, L., & Zheng, Y. (2019). Self-rated Attractiveness and Sociosexual Behavior Predict Gay Men's Preferences for Masculine Cues in Male Voices in China. *Evolutionary Psychology*, *17*(2), 147470491984743. https://doi.org/10.1177/1474704919847430

CHAPTER 3: TOPIC ANALYSIS REVEALS FIRST IMPRESSIONS OF VOICES

Ostrega, J., & Feinberg, D. R. (2023). Topic Analysis Reveals First Impressions of Voices.

PREFACE

After not being able to replicate the perceptual adaptation findings of normality judgements (Little et al., 2013) on attractiveness judgments, we began to consider the intricacies of voice perceptions and how those influence judgements of traits such as normality and attractiveness. Using the model of face perception and the perceptual face space (Oosterhof & Todorov, 2008), it became important for us to understand the complexities of the social perceptual voice space.

Some work had begun to test an understanding of perceptual voice space using rated traits similar to Oosterhof & Todorov (2008)'s work (McAleer et al., 2014). Using free form responses on a smaller sample of voices and faces, research unravelled some of the detailed understanding of people that we can gain from voices (Lavan, 2023). We expanded on both evaluations to improve our understanding further.

After participants were asked to provide free form responses of 100 speakers saying the word 'hi', using the aid of machine learning algorithms we were able to organize the information in these vast descriptions. We found results consistent with Lavan (2023)'s and that highlight the importance of valence and dominance found in Oosterhof & Todorov (2008) and replicated by McAleer et al., (2014). The hope is for

future work to build on these findings and provide a clearer understanding of voice impression and how they compare to the assumptions we form from faces.

REFERENCES

Lavan, N. (2023). How do we describe other people from voices and faces? *Cognition*, *230*, 105253. https://doi.org/10.1016/j.cognition.2022.105253

Little, A. C., Feinberg, D. R., DeBruine, L. M., & Jones, B. C. (2013). Adaptation to Faces and Voices: Unimodal, Cross-Modal, and Sex-Specific Effects. *Psychological Science*, *24*(11), 2297–2305. https://doi.org/10.1177/0956797613493293

McAleer, P., Todorov, A., & Belin, P. (2014). How Do You Say 'Hello'? Personality Impressions from Brief Novel Voices. *PLoS ONE*, *9*(3), e90779. https://doi.org/10.1371/journal.pone.0090779

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

CHAPTER 3: TOPIC ANALYSIS REVEALS FIRST IMPRESSIONS OF VOICES

ABSTRACT

People form rapid first impressions when encountering novel faces or voices. A popular theory on first impressions suggests there are two dimensions on which voices and faces vary: valence and dominance. Studies using orthogonal factor rotation, consistently find a third dimension in face space. However, participants are always given the same questionnaire. Unconstrained descriptions of female voices and faces from free-form responses have been tested but rely on researchers' opinions to categorize descriptions. To shift researcher degrees of freedom away from category membership, we used machine learning to categorize free responses to 50 female and 50 male voices saying 'hi'. We found a rich set of features that people use to categorize voices.

DATA AVAILABILITY STATEMENT

The datasets generated and analysed during the current study are available in the [Open Science Framework] repository, [https://osf.io/5t4b7/]

INTRODUCTION

People gather an incredible amount of information from exposure to novel voices and faces (Lavan, 2023). They form impressions about gender, sex, age, emotion, accent, socioeconomic status and many other traits from these stimuli (Collins, 2000; Feinberg et al., 2005, 2006; McAleer et al., 2014; Zuckerman et al., 1985). Whether or not this information is accurate, people use this information to form impressions.

A two-dimensional "social voice space" has been found for personality judgements of voices (Baumann & Belin, 2010; Latinus & Belin, 2011; Lavan et al., 2019; Lee et al., 2019). Acoustic components of voices have also been organized into a perceptual voice space. This voice space contains the two dimension of voice pitch and formant frequencies or vocal tract length (Baumann & Belin, 2010; Latinus & Belin, 2011; Lavan et al., 2019).

To explore impressions of voices and faces, including and beyond personality judgements, Lavan (2023) asked participants to provide unconstrained descriptions of women's faces and voices to examine if people form categories of faces and voices that are consistent with earlier findings. Physical characteristics, such as age and sex of a person were mentioned earlier than other characteristics. Psychological traits (personality characteristics) were mentioned most frequently for both voice and face descriptions. Although Lavan (2023) only used six voice stimuli, these findings show that we can draw information from voice descriptions, particularly that personality characteristics are not the only descriptors that occur when people make first impression.

We examined what impressions people form from novel voices and used topic analysis methods from machine learning to create categories. We predicted that we would find similar categories to those reported in Lavan (2023) and other work derived from theory-based rating sets (McAleer et al., 2014; Shiramizu et al., 2022).

METHODS

Stimuli

Voice clips of 100 individuals (50 female and 50 male) speaking the common English word “hi” were recorded in an anechoic sound-controlled booth (WhisperRoom Inc. SE 2000 Series Sound Isolation Enclosure), with speakers standing approximately 5–10 cm from the Sennheiser MKH 800 studio condenser microphone with a cardioid pick-up pattern. An M-Audio Fast Track Ultra interface was used to digitally encode the audio at a 96 kHz sampling rate and 32-bit amplitude quantization. Files were stored onto a computer as PCM WAV files using Adobe Soundbooth CS5 version 3.0. Previous work has shown that we can form first impressions from such short voice sounds such as “Hello” (McAleer et al., 2014). Voices were resampled to 44.1kHz and amplitude was normalized to 70dB RMS.

Procedure

Protocols for this study were approved by the McMaster University Research Ethics Board protocol #2296, all methods were performed in accordance with the relevant guidelines and regulations. Prior to analyses, we excluded all incomplete surveys. Participants were recruited from the McMaster University on-line subject pool (SONA) and were compensated with extra course credit or cash payment ($10 per hour pro rata)

for participation. The informed consent was obtained from all the participants of the study. There were 357 participants who completed this study.

After completing the consent from, participants were directed to a questionnaire in LimeSurvey where they were instructed to “Listen to the following voice and in the box provided please describe in as much detail as possible your impression of the speaker.” Underneath this instruction participants could play the voice clip using a sound player embedded in the question. They described their impression of each speaker, using a minimum of 200 characters, in a text box underneath. Playback repetitions were *ad libitum* and unmonitored.

All 100 voices were presented in a randomized order in a single block. Participants could not progress to the next voice until they have written at least 200 characters of their impression on the previous voice. Once all 100 voices were described participants were directed to a brief questionnaire regarding the participants self-report of sexual-orientation, sex, gender, age, whether English was their first language. We collected this data for demographic purposes but did not analyze our data with reference to it.

METHODS AND RESULTS

Data Cleaning

Participant data was compiled using Python version 3.10.0 (Van Rossum & Drake, 2009). Data was edited in Microsoft Excel (Microsoft Corporation, 2018) and any participants who did not describe all 100 voices or provided unusable data (i.e., copied and pasted descriptions as well as descriptions that did not reflect the task) were removed

from the final data set. Once all descriptions were compiled into a single list, we cleaned the data. We removed stopwords found in GENSIM's (Rehurek & Sojka, 2011) stopwords list, as well as adding our own list of stopwords found by running multiple iterations of this procedure. A full list of stopwords removed from the data can be found in our supplementary materials page here: https://osf.io/5t4b7/. We used Python Package Autocorrect (*Autocorrect 2.6.1*, 2021) to correct spelling mistakes. Duplicate descriptions were removed using the *Pandas drop_duplicates* method. Words less than or equal to 3 characters long, and numeric values were also removed from the data. In total there were 34,100 descriptions containing a total of 468,695 words.

LDA Methods

We used Latent Dirichlet Allocation to analyze the topics of texts. To prepare the text for analysis, we first dropped any duplicate response. Then we removed a list of stopwords that included those found in GENSIM (Rehurek & Sojka, 2011), and a list of other high frequency words irrelevant to the analysis (see supplementary material for complete list of added stopwords). We removed any words less than 3 letters, punctuation, and then stemmed the words. Then we used GENSIM to generate bigrams and trigrams. Subsequently we split each description into a list of words and created a bag of words corpus. We used the LDA function in GENSIM to generate models using the maximum coherence value as a hyperparameter to choose the optimal number of topics (between 2 and 25, inclusive). The coherence score represents how similar topic keywords are to each other and is used as a proxy to pick the best model.

LDA Results

We found that 13 topics yielded the highest coherence score which was 10. This score was relatively low, and topics contained multiple concepts, and some concepts spanned multiple topics. Because of this we did a secondary analysis of the topic keywords using ChatGPT 4.0 (OpenAI, 2023) to categorize their semantic field.

Semantic field methods

To generate categories of words to use with a large language model, we first took all the top 10 keywords from the top 10 topics generated by the LDA model, then asked Chat GPT 4.0 (OpenAI, 2023) using a series of queries available in the supplementary materials, to generate those categories.

Semantic field results

Category	Number of LDA words in each category	Frequency of LDA words in category
Age-Related Descriptors	28	30.769
Valence	15	16.484
Gender & Sexual Orientation	11	12.088
Linguistic & Accents	9	9.890
Social Context & Interactions	9	9.890
Dominance	8	8.791
Educational & Professional Context	3	3.297
Media & Entertainment	3	3.297
Miscellaneous	3	3.297
Physical Appearance	2	2.198

TABLE 1 This table shows the categories and number of keywords within each category.

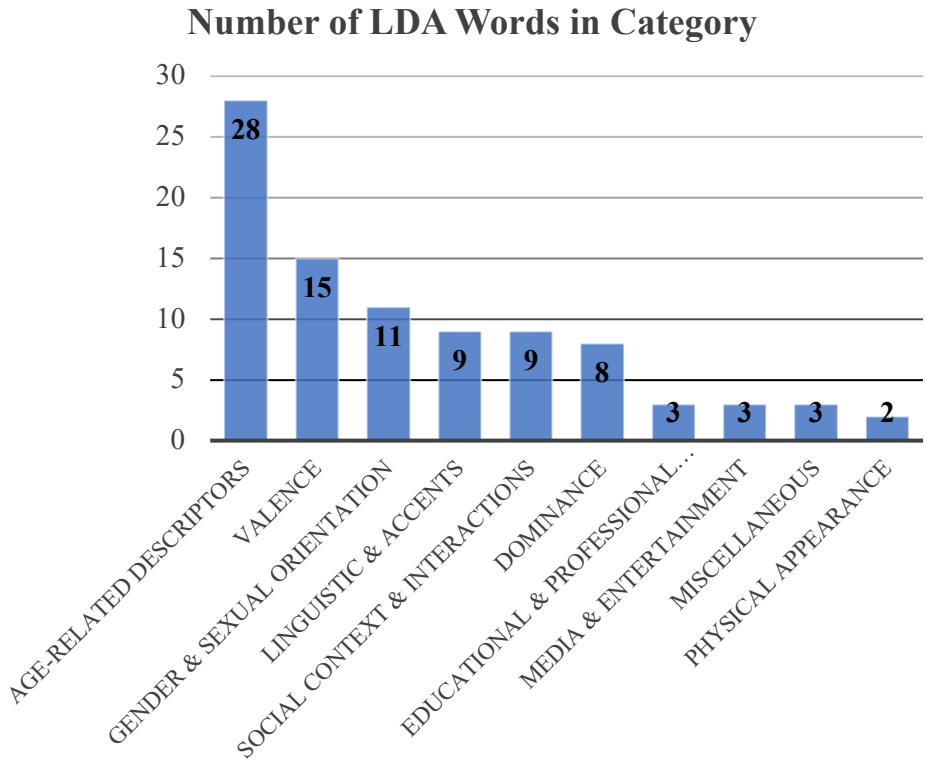

FIGURE 1 This graph shows the categories and number of keywords within each category.

DISCUSSION

We found that the human voice conveys a rich set of attributes. When forming first impressions of voices, people in our study formed impressions of age, valence, gender & sexual orientation, Linguistics & Accents, Dominance, Education & Professional Context, Media & Entertainment, & Physical Appearance. (in order of topic keyword frequency). It is likely that other studies using different stimuli and raters in different places and ages will find different categories to be more or less important (e.g., children may be less likely to think about terms related to mating than would adults).

Dominance and valence were important categories as expected (Oosterhof & Todorov, 2008). However, people talked the most about how old people sounded.

About one-third of the topic keywords were summarized using valence and dominance. This highlights the importance of valence and dominance in personality impressions from voices ((McAleer et al., 2014; Oosterhof & Todorov, 2008). While personality judgements are significant in impression formation, the intension of our results were to explore all person perceptions from voices, including and beyond personality variables.

Our participants used a wide range of descriptors when asked to describe voices using free form responses, which is very similar to Lavan (2023). Participants described characteristics of voicing using personality, physical, social, and auditory traits. Congruent with other work on voice impressions (Collova et al., 2019; Lavan, 2023; Oosterhof & Todorov, 2008; Sutherland et al., 2013) many descriptors were of personality attributes. Physical traits were mentioned the most, followed by psychological traits, social traits, and voice quality.

There was a salient presence of physical descriptions, specifically of age and sex occurring early in participant descriptions and consistent with Lavan (2023) findings. Social, auditory, and other traits were mentioned and were central to some of the topics generated by our model.

Our study demonstrates the strength of using free form description data to learn about how people perceive voices (Lavan, 2023; Mondloch et al., 2022). Free descriptions allow for participants to focus on the aspects of voices and other stimuli they

find important and provide a fuller account of their impression without being restricted by attributes presented by the experimenter (Lavan, 2023; Mondloch et al., 2022). Free descriptions were used in prior work (Lavan, 2023; Oosterhof & Todorov, 2008), and our machine-learning techniques add support to these studies.

REFERENCES

Autocorrect 2.6.1. (2021). https://pypi.org/project/autocorrect/

Baumann, O., & Belin, P. (2010). Perceptual scaling of voice identity: Common dimensions for different vowels and speakers. *Psychological Research Psychologische Forschung*, *74*(1), 110–120. https://doi.org/10.1007/s00426-008-0185-z

Collins, S. A. (2000). Men's voices and women's choices. *Animal Behaviour*, *60*(6), 773–780. https://doi.org/10.1006/anbe.2000.1523

Collova, J. R., Sutherland, C. A. M., & Rhodes, G. (2019). Testing the functional basis of first impressions: Dimensions for children's faces are not the same as for adults' faces. *Journal of Personality and Social Psychology*, *117*(5), 900–924. https://doi.org/10.1037/pspa0000167

Feinberg, D. R., Jones, B. C., Law Smith, M. J., Moore, F. R., DeBruine, L. M., Cornwell, R. E., Hillier, S. G., & Perrett, D. I. (2006). Menstrual cycle, trait estrogen level, and masculinity preferences in the human voice. *Hormones and Behavior*, *49*(2), 215–222. https://doi.org/10.1016/j.yhbeh.2005.07.004

Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour*, *69*(3), 561–568. https://doi.org/10.1016/j.anbehav.2004.06.012

Latinus, M., & Belin, P. (2011). Anti-Voice Adaptation Suggests Prototype-Based Coding

of Voice Identity. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00175

Lavan, N. (2023). How do we describe other people from voices and faces? *Cognition*, *230*, 105253. https://doi.org/10.1016/j.cognition.2022.105253

Lavan, N., Knight, S., & McGettigan, C. (2019). Listeners form average-based representations of individual voice identities. *Nature Communications*, *10*(1), 2404. https://doi.org/10.1038/s41467-019-10295-w

Lee, Y., Keating, P., & Kreiman, J. (2019). Acoustic voice variation within and between speakers. *The Journal of the Acoustical Society of America*, *146*(3), 1568–1579. https://doi.org/10.1121/1.5125134

McAleer, P., Todorov, A., & Belin, P. (2014). How Do You Say 'Hello'? Personality Impressions from Brief Novel Voices. *PLoS ONE*, *9*(3), e90779. https://doi.org/10.1371/journal.pone.0090779

Microsoft Corporation. (2018). *Microsoft Excel* [Computer software]. https://office.microsoft.com/excel

Mondloch, C. J., Twele, A. C., & Thierry, S. M. (2022). We need to move beyond rating scales, white faces and adult perceivers: Invited Commentary on Sutherland & Young (2022), understanding trait impressions from faces. *British Journal of Psychology*, bjop.12619. https://doi.org/10.1111/bjop.12619

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

OpenAI. (2023). ChatGPT (Version 4.0) [Language model]. Retrieved from https://www.example.com/chatgpt4

Rehurek, R., & Sojka, P. (2011). *Gensim–python framework for vector space modelling*. *3*(2).

Shiramizu, V. K. M., Lee, A. J., Altenburg, D., Feinberg, D. R., & Jones, B. C. (2022). The role of valence, dominance, and pitch in perceptions of artificial intelligence (AI) conversational agents' voices. *Scientific Reports*, *12*(1), 22479. https://doi.org/10.1038/s41598-022-27124-8

Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Michael Burt, D., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition*, *127*(1), 105–118. https://doi.org/10.1016/j.cognition.2012.12.001

Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace.

Zuckerman, M., Driver, R., & Guadagno, N. S. (1985). Effects of segmentation patterns on the perception of deception. *Journal of Nonverbal Behavior*, *9*(3), 160–168. https://doi.org/10.1007/BF01000737

CHAPTER 4: ARE YOU LISTENING? PERCEPTIONS OF VOICES LABELLED AS DOCTORS AND NURSE ON TRUSTWORTHINESS, COMPETENCE, DOMINANCE, AND ATTRACTIVENESS JUDGEMENTS

PREFACE

From our work on voice attraction perception and evaluating the impressions that people form from simply hearing someone speak the word ‘hi’, we wanted to examine how voice impressions can be applied to a novel setting using characteristics that have been critical to understanding voices and faces in previous literature. It is becoming more common in the healthcare sector to communicate with care providers through voice-based communications like over the telephone. This is why we felt it vital to examine voice impressions in this setting specifically, especially considering that these impressions can play a critical role in care provision (Liu et al., 2020; Schattner et al., 2004; von Bültzingslöwen et al., 2006).

The two key dimensions of the social face and voice space models are trustworthiness and dominance (McAleer et al., 2014; Oosterhof & Todorov, 2008), this is why we had participants provide judgements of voices on these two attributes. We also asked participants to judge voices on attractiveness because of the importance it plays in our social heuristics of others (Dion et al., 1972), but also our understanding of how critical aspects of voice pitch influences these judgements and to see if we once again could replicate these findings in a novel setting. We also evaluated competence judgements because of the importance competence evaluations play in a professional and

leadership context, but also its relationship to dominance judgements (Anderson & Klofstad, 2012; Cartei et al., 2021; Klofstad et al., 2012; Rogers et al., 2020; Rotter & Stein, 1971).

We found that when labelled as doctors, low-pitched female voices were judged as more competent than male voices. Competence ratings were similar for both sexes when they were labeled as belonging to nurses. Low-pitched voices were judged as more dominant regardless of voice sex and profession. High-pitched female voices and low-pitched male voices were judged as most attractive regardless of profession.

REFERENCES

Anderson, R. C., & Klofstad, C. A. (2012). Preference for Leaders with Masculine Voices Holds in the Case of Feminine Leadership Roles. *PLoS ONE*, *7*(12), e51216. https://doi.org/10.1371/journal.pone.0051216

Cartei, V., Oakhill, J., Garnham, A., Banerjee, R., & Reby, D. (2021). Voice Cues Influence Children's Assessment of Adults' Occupational Competence. *Journal of Nonverbal Behavior*, *45*(2), 281–296. https://doi.org/10.1007/s10919-020-00354-y

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, *24*(3), 285–290. https://doi.org/10.1037/h0033731

Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: Voice pitch

influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1738), 2698–2704. https://doi.org/10.1098/rspb.2012.0311

Liu, S., Zhang, M., Gao, B., & Jiang, G. (2020). Physician voice characteristics and patient satisfaction in online health consultation. *Information & Management*, *57*(5), 103233. https://doi.org/10.1016/j.im.2019.103233

McAleer, P., Todorov, A., & Belin, P. (2014). How Do You Say 'Hello'? Personality Impressions from Brief Novel Voices. *PLoS ONE*, *9*(3), e90779. https://doi.org/10.1371/journal.pone.0090779

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Rogers, K., Bryant, D., & Howard, A. (2020). Robot Gendering: Influences on Trust, Occupational Competency, and Preference of Robot Over Human. *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–7. https://doi.org/10.1145/3334480.3382930

Rotter, J. B., & Stein, D. K. (1971). Public Attitudes Toward the Trustworthiness, Competence, and Altruism of Twenty Selected Occupations. *Journal of Applied Social Psychology*, *1*(4), 334–343. https://doi.org/10.1111/j.1559-1816.1971.tb00371.x

Schattner, A., Rudin, D., & Jellin, N. (2004). Good physicians from the perspective of

their patients. *BMC Health Services Research*, *4*(1), 26. https://doi.org/10.1186/1472-6963-4-26

von Bültzingslöwen, I., Eliasson, G., Sarvimäki, A., Mattsson, B., & Hjortdahl, P. (2006). Patients' views on interpersonal continuity in primary care: A sense of security based on four core foundations. *Family Practice*, *23*(2), 210–219. https://doi.org/10.1093/fampra/cmi103

CHAPTER 4: ARE YOU LISTENING? PERCEPTIONS OF VOICES LABELLED AS DOCTORS AND NURSE ON TRUSTWORTHINESS, COMPETENCE, DOMINANCE, AND ATTRACTIVENESS JUDGEMENTS

ABSTRACT

Since the COVID-19 pandemic, remote care provision through telephone and online communication has become more prevalent, making voice-based impressions more important. Valence and dominance are important perceptual dimensions that could impact patient care. We investigated how the characteristics of doctors' and nurses' voices influence perceptions of competence, trustworthiness, dominance, and attractiveness, based on brief first impressions. Participants rated 50 male or 50 female voices on the aforementioned traits and were told the people were either doctors or nurses. We found that voice pitch effects perceptions of doctors and nurses differently. Female raters judged nurses as more dominant than did male raters. Low pitched voices were rated as more dominant sounding than high pitched voices, we observed effects in the same directions for competence judgements, but these effects were not statistically significant. Female doctors were judged as more competent sounding than male doctors. High-pitched female and low-pitched male voices were judged as more attractive across both professions. The same pattern of results was observed for trustworthiness ratings but were not statistically significant. Stereotypically gender atypical professions seem to reduce the association between voice pitch and the traits examined.

INTRODUCTION

Since the COVID-19 pandemic, there has been in increase in care provision by physicians and nurses through telephone appointments. People tend to rate professionals like doctors highly on trustworthiness, competence, and altruism (Rotter & Stein, 1971). When patients do not trust their healthcare provider they are less likely to seek medical care, engage in preventative care and follow recommendation from their healthcare provider (Schattner et al., 2004; von Bültzingslöwen et al., 2006). Since trust is inferred from the voice, particularly in the absence of other cues (McAleer et al., 2014), the way a doctor or nurse sounds could have effects on patient care.

Dominance, valence, competence, and attractiveness explain the majority of variance in voice perception (McAleer et al., 2014). Voice pitch has been shown to predict perceptions of dominance, but not valence (trust and attractiveness) of artificial intelligent (AI) conversational agents such as Siri and Alexa (Shiramizu et al., 2022). Voice pitch is also implicated in selection of political leaders (Anderson & Klofstad, 2012; Klofstad et al., 2012, 2015; Schild et al., 2022). Blind and sighted people use voice pitch the same way to judge competence, trust, and warmth, and cues to dominance and attractiveness as do sighted people (Oleszkiewicz et al., 2017; Pisanski et al., 2017). Stereotypical voice-based judgements of occupational competence already present in children that parallel those found in adults (Cartei et al., 2021). Since these perceptions of the voice appear in multiple domains and do not require visual input, and are present from a very early age, it is likely that these associations will also extend to people's perceptions of doctors and nurses. However, it is possible that gender stereotypes associated with male nurses and

female doctors could qualify the relationship between voice pitch and the aforementioned perceptual dimensions.

For example, patients seem to be more satisfied with female doctors who behave in line with stereotypically female behaviour (Zelek & Phillips, 2003). Male physicians are received more positively when using aggressive communication styles, but female physicians were perceived more negatively when using aggressive communication styles (Bradley et al., 2001; Burgoon et al., 1991).

Mast et al., (2008) showed that voice attractiveness positively predicted patient satisfaction. Patient satisfaction was related to doctors having a male-gender stereotype of having a louder voice. This suggests that low-pitched voices that sound dominant and attractive may also elicit higher patient satisfaction, which could in-turn affect patient care.

High-pitched women's voice and low-pitched men's voices are generally perceived to be relatively more attractive (Apicella & Feinberg, 2009; Collins, 2000; Feinberg et al., 2005, 2008; Jones et al., 2008, 2010; Puts et al., 2006; Re et al., 2012; Vukovic et al., 2008; Zhang et al., 2018). The "what sounds beautiful is good" stereotype also suggests that people ascribe positive qualities to attractive voices. (Zuckerman & Driver, 1988). Interestingly, listeners rated male speakers with high-pitched voices high on attractiveness when they spoke antisocial words compared to male speakers with low-pitched voices, and consistently rated male voices speaking prosocial words as more attractive regardless of voice pitch (O'Connor & Barclay, 2018). Collectively, these studies suggest that voice pitch could affect our perceptions of doctors and nurses.

Perceptions of dominance are also strongly negatively associated with pitch (Armstrong et al., 2019) among male and female vocalizers. Surgeons with more dominant sounding voices were more likely to have a history of malpractice claims than those with less dominant sounding voices (Ambady et al., 2002). In this case, a low voice pitch might be perceived as a negative trait for doctors and nurses.

The influence of voice pitch on trustworthiness perceptions is equivocal. Some studies find that people tend to trust higher pitched male voices more than lower pitched male voice both generally (McAleer, Todorov, & Belin, 2014; O'Connor & Barclay, 2017) and in financial contexts (Montano, Tigue, Isenstein, Barclay, & Feinberg, 2017; O'Connor & Barclay, 2017). Other results show that lower pitched voices are generally perceived are more trustworthy than higher pitched voices in both male (Oleszkiewicz, Pisanski, Lachowicz-Tabaczek, & Sorokowska, 2017; Tigue et al., 2012) and female speakers (Klofstad, Anderson, & Peters, 2012). Furthermore, other work has not detected any significant effects of voice pitch on perceptions of trustworthiness (Klofstad, Anderson, & Peters, 2012; Vukovic et al., 2011). Although low-pitched male voices and high-pitched female voices correlate with attractiveness judgments (reviewed in Feinberg, 2008; Puts, Jones, & DeBruine, 2012), people generally perceive lower pitched male voices and higher pitched female voices as sounding untrustworthy in romantic contexts (O'Connor & Feinberg, 2012; O'Connor, Pisanski, Tigue, Fraccaro, & Feinberg, 2014; O'Connor, Re, & Feinberg, 2011). Accordingly, voice pitch could influence how trustworthy we judge doctors and nurses to be.

We therefore explored the role of voice pitch in perceptions of trust, dominance, competence, and attractiveness of people's voices. We hypothesized that:

Voice pitch is strongly and negatively related to perceptions of dominance and competence (Armstrong et al., 2019; Borkowska & Pawlowski, 2011; Hodges-Simeon et al., 2010; Klofstad et al., 2015; Puts et al., 2006; Schild et al., 2022). We predicted that regardless of gender and profession, people with low pitched voices would be judged as more dominant and competent. We also predicted that low pitched men's voices and high pitch in women's voices would be judged as relatively more trustworthy and attractive than high-pitched men's voices and low-pitched women's voices.

Given prior evidence of gender stereotypes in perceptions of doctors and nurses (Bradley et al., 2001; Burgoon et al., 1991; Cartei et al., 2021; Dielissen et al., 2011; Mast et al., 2008; Stephens et al., 2016; Zelek & Phillips, 2003), we also expected that men would be perceived as more competent, dominant, attractive, and trustworthy doctors than women would be judged as. We also predicted that women would be perceived as more competent, dominant, attractive, and trustworthy nurses than men would be judged as. Aggressiveness also affect's perception competence of male doctors more than it affects perceptions of competence from female doctors (Bradley et al., 2001; Burgoon et al., 1991). Voice pitch is negatively tied to perceptions of aggressiveness (Aung & Puts, 2020; Zhang & Reid, 2017). Therefore, we predicted that voice pitch may affect competence ratings of male doctors more than female doctors and female or male nurses.

METHODS

Stimuli

Previous work has shown that we can form first impressions from monosyllabic voice sounds (Bruckert et al., 2010; Pisanski, Groyecka-Bernard, & Sorokowski, 2021). We used voice clips of 100 individuals, 50 female (mean f0 range = 147.23 – 288.70 Hz) and 50 male (mean f0 range = 71.13 – 186.43 Hz) voices speaking the common English word “hi”, aiming to increase the size of the stimuli set used by McAleer et al. (2014). These voices were recorded in an anechoic sound-controlled booth (WhisperRoom Inc. SE 2000 Series Sound Isolation Enclosure), with vocalizers standing approximately 5–10 cm from the Sennheiser MKH 800 studio condenser microphone with a cardioid pick-up pattern. An M-Audio Fast Track Ultra interface was used to digitally encode the audio at a 96 kHz sampling rate and 24-bit amplitude quantization. Files were stored onto a computer as WAV. files using Adobe Soundbooth CS5 version 3.0. Voices were resampled to 44.1kHz and amplitude was normalized to 70dB RMS.

Participants

There were 2359 participants; 1309 men (mean age = 34.34), 1002 women (mean age = 35.58), 36 non-binary (mean age = 26.53), 2 two-spirited (mean age = 21), 6 preferred not to answer (mean age = 30), and 4 others (mean age = 26.25) who completed the study. There were approximately 150 participants in each of the 16 conditions, exact values specified in TABLE 1. Participants were randomly assigned to conditions. Each participant rated 50 voices of one sex, one profession, and on one trait. The 16 conditions included all possible combinations of: 4 ratings (dominance, attractiveness, competence,

trust); 2 sexes of voice (female, male); and 2 professions (doctor, nurse). All conditions were between-subject.

	Doctor		**Nurse**	
Rated Trait	**Female Voice**	**Male Voice**	**Female Voice**	**Male Voice**
Competence	144	162	154	158
Attractiveness	167	120	155	139
Dominance	147	164	140	140
Trust	141	131	136	161

TABLE 1 The number of participants who rated voices across all 16 conditions. With each person rating 50 voices of one sex, one profession, on one trait. The 16 conditions included all possible combinations of: 4 ratings (dominance, attractiveness, competence, trust); 2 sexes of voice (female, male); and 2 professions (doctor, nurse).

Procedure

Protocols for this study were approved by the McMaster University Research Ethics Board protocol #6295, all methods were performed in accordance with the relevant guidelines and regulations. Prior to analyses, we excluded all incomplete surveys. Participants were recruited from the Prolific.co on-line subject pool and were compensated 4.76 GBP (8 CAD) for 30 minutes of participation. The informed consent was obtained from all the participants of the study.

Participants were then presented with voice labelled as belonging to a nurse or a doctor. They rated on a scale from 1-7 how competent, trustworthy, dominant, or attractive each voice sounded to them. Where a rating of 1 represented a voice sounding low on the trait and a rating of 7 being high on the trait. If the trait was attractiveness a rating of 1 indicated the voice sounded very unattractive and a rating of 7 indicated the voice sounded very attractive. The voices were presented in a randomized order. Each participant rated 50 voices for one trait for one sex of voice for one profession. Each trial

indicated to the participant the profession of the speaker and what trait they were rating the speaker on.

Each participant completed a short demographic questionnaire about age, sex, gender, and sexual orientation after rating the voices. The experiment concluded with participants being asked to re-consent after being debriefed about the study and the deception that the voices did not belong to nurses or doctors.

RESULTS

We compiled our data using R version 4.3.1 (R Core Team, 2021), we organized our data using tidyverse 1.3.1 (Wickham et al., 2019) and used multilevel linear regression with lmerTest 3.1.3 (Kuznetsova, A., Brockhoff, P. B., & Christensen R. H. B., 2017) and jtools 2.1.4 (Long, 2020) to analyze our data. We conducted separate analyses for each rating. We also effect coded our response variables by z-scoring them to have a mean of 0 and a standard deviation of 1. This was done to ensure that the values were on comparable scales, and that the impact of the variables were easily understandable.

In each separate analysis, the response variable was the rating type (competence, dominance, attractiveness, or trustworthiness). The fixed effects were profession (doctor or nurse), voice sex (female or male), and rater sex (female or male). We entered random intercepts for each rater (participant) and each voice (sound file), and random slopes for each sex of rater for each voice. There were no-within subjects' comparisons.

Competence

There was a three-way interaction between voice pitch, the sex of the voice and profession, a two-way interaction of voice pitch and profession and main effect of

profession on competence judgments. There was a negative correlation between voice pitch and competence perceptions. Lower pitched voices were judged to be more competent sounding then higher pitched voices. This negative relationship was stronger when participants rated male voices compared to female voices, and female voices were judged as more competent sounding, but only when the voices were labelled as belonging to a doctor. There was no difference between competence ratings for male and female speakers when participants rated nurses. Nurses were generally perceived as more competent sounding than doctors. These results are depicted below in FIGURE 1.

```
competence_model <- lmer(z_score_rating_value ~ Pitch_z_score * sex_of_voice_effect_coded
* rater_sex_effect_coded * profession_effect_coded +
         (1 | participant) +
         (1 + rater_sex_effect_coded * profession_effect_coded * study | sound_file),
         REML = FALSE,
         control = lmerControl(optimizer ="bobyqa", optCtrl=list(maxfun=100000)),
         data=competence_data)
```

	unstandardized estimate	**standard error**	**t value**	**d.f.**	**p**
Intercept	0.150	0.038	3.947	187.095	<0.0001***
Voice Pitch (z-scored)	-0.059	0.031	-1.935	100.457	0.056
Sex of Voice (effect coded)	-0.096	0.076	-1.260	187.095	0.209
Rater Sex (effect coded)	-0.055	0.042	-1.320	634.973	0.187
Profession (effect coded)	0.083	0.041	2.022	619.443	0.044*
Voice Pitch x Sex of Voice	-0.025	0.061	-0.412	100.457	0.681
Voice Pitch x Rater Sex	0.005	0.012	0.390	117.211	0.697
Sex of Voice x Rater Sex	-0.027	0.083	-0.320	634.973	0.749
Voice Pitch x Profession	0.046	0.009	4.910	377.883	<0.0001***
Sex of Voice x Profession	0.105	0.082	1.285	619.443	0.199
Rater Sex x Profession	0.075	0.082	0.916	620.326	0.360
Voice Pitch x Sex of Voice x Rater Sex	-0.014	0.023	-0.603	117.211	0.548
Voice Pitch x Sex of Voice x Profession	0.053	0.019	2.803	377.883	0.005**
Voice Pitch x Rater Sex x Profession	-0.020	0.020	-1.024	197.886	0.307
Sex of Voice x Rater Sex x Profession	-0.150	0.165	-0.912	620.326	0.362
Voice Pitch x Sex of Voice x Rater Sex x Profession	-0.064	0.039	-1.646	197.886	0.101

TABLE 2 Results of our analysis for competence ratings by speaker voice pitch (z-scored), sex of the speaker (effect coded), rater sex (effect coded), and speaker profession (effect coded), with a confidence interval of 95%. *p<0.05; **p<=0.01; ***p<=0.0001

FIGURE 1 Scatter plot of our analysis for competence ratings by speaker voice pitch (z-scored), sex of the speaker, and speaker profession.

Attractiveness

With attractiveness judgements there was a four-way interaction between voice pitch, sex of the voice, sex of the voice and rater sex. A moderate positive relationship between voice pitch and attractiveness judgments was observed when participants rated female voices; high pitched female voices were judged to be more attractive than low

pitched female voice. Conversely, a stronger negative relationship was found between voice pitch and attractiveness judgements when participants rated male voices. Lower pitched voices were rated as more attractive than high pitched voices for male speakers (FIGURE 2). Male raters judged female speakers to be more attractive sounding than female raters did and showed a stronger positive correlation between voice pitch and attractiveness judgments in female speakers. Female raters showed a weaker positive correlation between voice pitch and attractiveness for female speakers (FIGURE 3). A very strong negative correlation was observed between voice pitch and attractiveness ratings when female raters judged male speakers, this correlation was weaker but still in the same direction for male raters (FIGURE 3).

A two-way interaction was found between voice pitch and rater sex. High-pitched voices were judged as more attractive sounding for female speakers and low-pitched voices were judged as more attractive for male speakers. Main effects were also found for the sex of the voice and voice pitch. Higher pitched and female voices were judged to be the most attractive.

```
attractiveness_model <- lmer(z_score_rating_value ~ Pitch_z_score * sex_of_voice_
effect_coded * rater_sex_effect_coded * profession_effect_coded +
          (1 | participant) +
          (1 + rater_sex_effect_coded * profession_effect_coded * study | sound_file)
          REML = FALSE,
          control = lmerControl(optimizer ="bobyqa", optCtrl=list(maxfun=100000)),
          data=attractiveness_data)
```

	unstandardized estimate	**standard error**	**t value**	**d.f.**	**p**
Intercept	-0.073	0.032	-2.285	266.099	0.023*
Voice Pitch (z-scored)	-0.070	0.023	-3.020	100.167	0.003**
Sex of Voice (effect coded)	-0.230	0.063	-3.633	261.602	<0.0001***
Rater Sex (effect coded)	0.007	0.043	0.163	593.313	0.871
Profession (effect coded)	0.048	0.043	1.129	578.955	0.259
Voice Pitch x Sex of Voice	-0.190	0.046	-4.079	100.167	<0.0001***
Voice Pitch x Rater Sex	0.032	0.012	2.635	108.227	0.010*
Sex of Voice x Rater Sex	-0.227	0.085	-2.680	633.736	0.008**
Voice Pitch x Profession	0.010	0.011	0.910	127.008	0.365
Sex of Voice x Profession	-0.128	0.084	-1.523	619.709	0.128
Rater Sex x Profession	-0.054	0.085	-0.630	578.265	0.529
Voice Pitch x Sex of Voice x Rater Sex	0.010	0.025	0.414	108.227	0.680
Voice Pitch x Sex of Voice x Profession	-0.017	0.021	-0.786	127.008	0.433
Voice Pitch x Rater Sex x Profession	-0.028	0.020	-1.444	546.341	0.149
Sex of Voice x Rater Sex x Profession	-0.102	0.167	-0.611	620.798	0.541
Voice Pitch x Sex of Voice x Rater Sex x Profession	0.112	0.039	2.861	546.341	0.004**

TABLE 3 Results of our analysis for attractiveness ratings by speaker voice pitch (z-scored), sex of the speaker (effect coded), rater sex (effect coded), and speaker profession (effect coded), with a confidence interval of 95%. *p<0.05; **p<=0.01; ***p<=0.0001

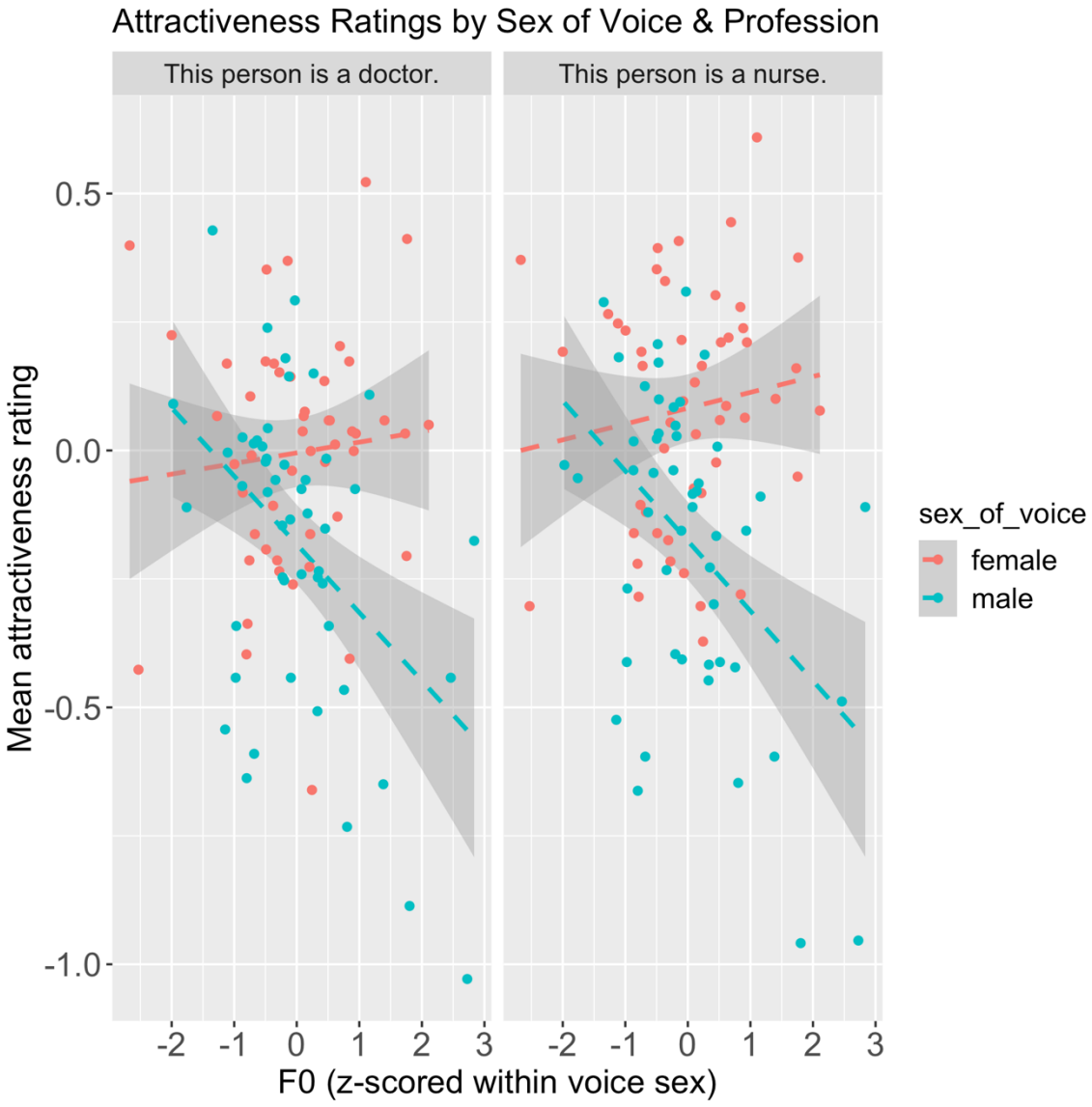

FIGURE 2 Scatter plot of our analysis for attractiveness ratings by speaker voice pitch (z-scored), sex of the speaker, and speaker profession.

FIGURE 3 Scatter plot of our analysis for attractiveness ratings by speaker voice pitch (z-scored), sex of the speaker, and rater sex.

Dominance

With dominance judgements there was a three-way interaction between voice pitch, sex of the rater, and profession. There was a strong negative relationship between voice pitch and dominance judgements. Low pitched voices were judged as more dominant sounding than high pitched voices. Female raters judged nurses as sounding

more dominant than did male raters, the same was not true for judgments of doctors (FIGURE 5). A two-way interaction between voice pitch and profession was found. Nurses were rated as more dominant sounding than doctors. A main effect of voice pitch was observed with lower pitched voices rated as more dominant sounding.

```
dominance_model <- lmer(z_score_rating_value ~ Pitch_z_score * sex_of_voice_effect_
coded * rater_sex_effect_coded * profession_effect_coded +
          (1 | participant) +
          (1 + rater_sex_effect_coded * profession_effect_coded * study | sound_file),
          REML = FALSE,
          control = lmerControl(optimizer ="bobyqa", optCtrl=list(maxfun=100000)),
          data=dominance_data)
```

	unstandardized estimate	**standard error**	**t value**	**d.f.**	**p**
Intercept	-0.231	0.038	-6.020	191.592	<0.0001***
Voice Pitch (z-scored)	-0.190	0.032	-6.031	100.213	<0.0001***
Sex of Voice (effect coded)	0.140	0.077	1.822	191.592	0.070
Rater Sex (effect coded)	-0.074	0.044	-1.691	615.897	0.091
Profession (effect coded)	0.037	0.043	0.873	592.761	0.383
Voice Pitch x Sex of Voice	-0.050	0.063	-0.789	100.213	0.432
Voice Pitch x Rater Sex	-0.007	0.014	-0.504	105.676	0.615
Sex of Voice x Rater Sex	-0.060	0.087	-0.693	615.897	0.488
Voice Pitch x Profession	-0.021	0.011	-1.990	183.857	0.048*
Sex of Voice x Profession	0.107	0.085	1.253	592.761	0.211
Rater Sex x Profession	-0.129	0.085	-1.516	594.797	0.130
Voice Pitch x Sex of Voice x Rater Sex	0.015	0.028	0.538	105.676	0.592
Voice Pitch x Sex of Voice x Profession	0.034	0.021	1.627	183.857	0.106
Voice Pitch x Rater Sex x Profession	0.061	0.022	2.796	150.373	0.006**
Sex of Voice x Rater Sex x Profession	0.250	0.171	1.462	594.797	0.144
Voice Pitch x Sex of Voice x Rater Sex x Profession	0.029	0.044	0.660	150.373	0.510

TABLE 4 Results of our analysis for dominance ratings by speaker voice pitch (z-scored), sex of the speaker (effect coded), rater sex (effect coded), and speaker profession (effect coded), with a confidence interval of 95%. *p<0.05; **p<=0.01; ***p<=0.0001

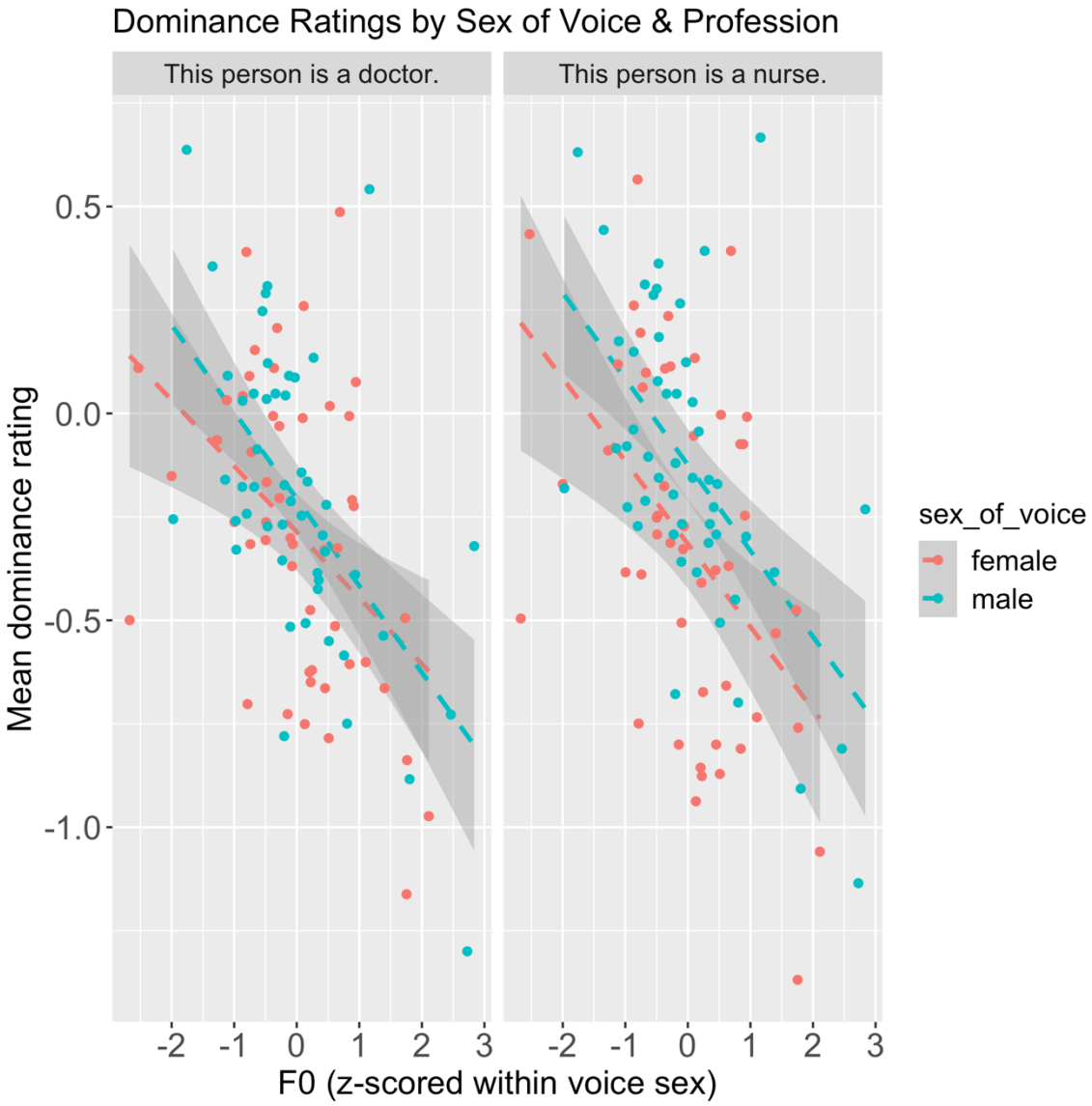

FIGURE 4 Scatter plot of our analysis for dominance ratings by speaker voice pitch (z-scored), sex of the speaker, and speaker profession.

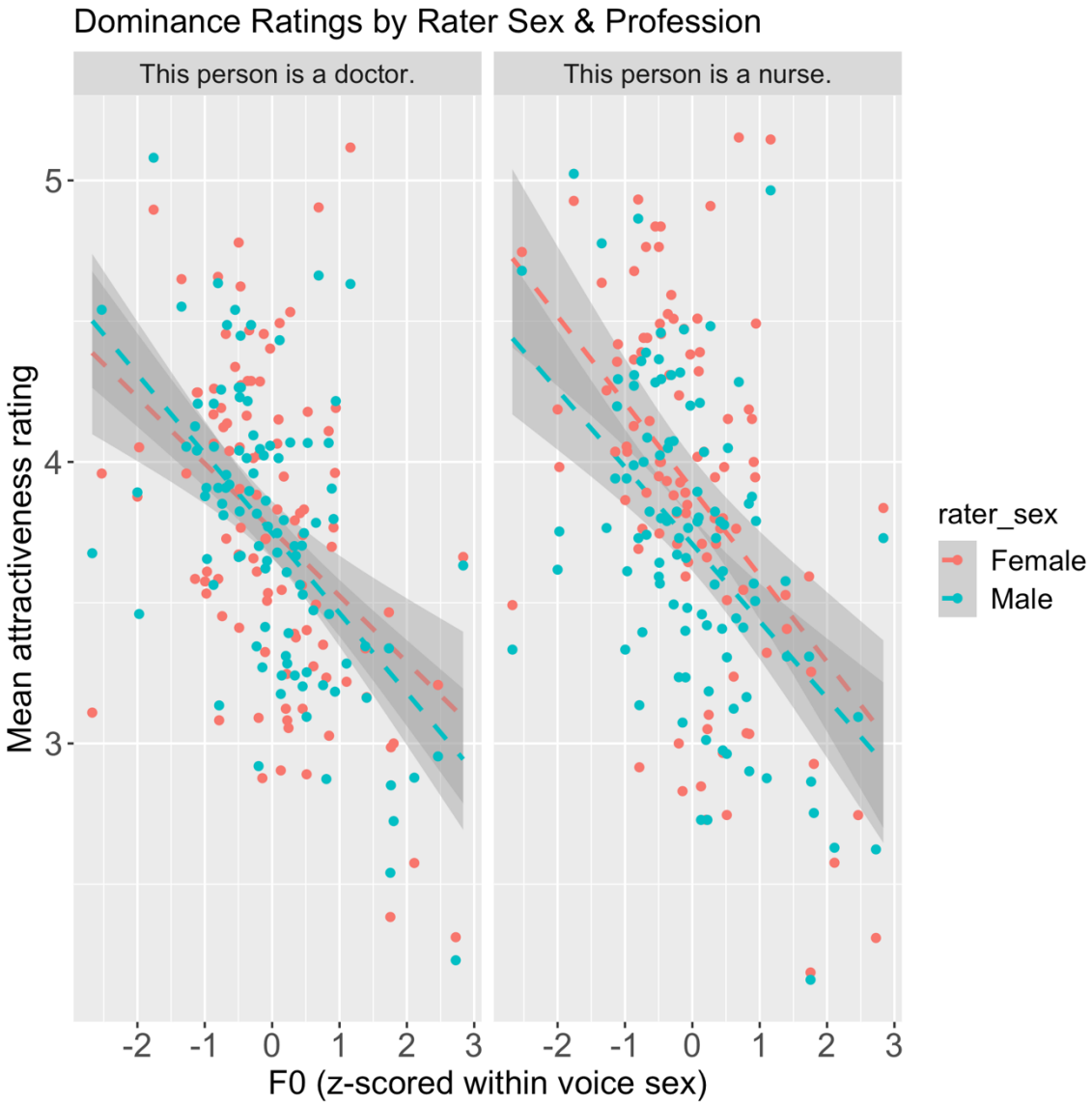

FIGURE 5 Scatter plot of our analysis for dominance ratings by speaker voice pitch (z-scored), sex of rater, and rater sex.

Trustworthiness

Although our findings trend towards a positive correlation for voice pitch on trustworthiness judgements for female speakers and a negative correlation for voice pitch on trustworthiness judgments for male speakers (FIGURE 6). There were no significant

main effects or interactions between speaker voice pitch, speaker sex, rater sex or profession.

```
trust_model <- lmer(z_score_rating_value ~ Pitch_z_score * sex_of_voice_effect_coded *
rater_sex_effect_coded * profession_effect_coded +
          (1 | participant) +
          (1 + rater_sex_effect_coded * profession_effect_coded * study | sound_file),
          REML = FALSE,
          control = lmerControl(optimizer ="bobyqa", optCtrl=list(maxfun=100000))
          data=trustworthiness_data)
```

	unstandardized estimate	**standard error**	**t value**	**d.f.**	**p**
Intercept	0.115	0.033	4.625	217.414	<0.0001***
Voice Pitch (z-scored)	-0.009	0.026	-0.362	99.843	0.718
Sex of Voice (effect coded)	-0.062	0.067	-0.920	217.414	0.358
Rater Sex (effect coded)	0.041	0.040	1.017	575.009	0.309
Profession (effect coded)	-0.051	0.040	-1.267	570.816	0.206
Voice Pitch x Sex of Voice	-0.048	0.052	-0.932	99.843	0.354
Voice Pitch x Rater Sex	0.017	0.010	1.611	173.570	0.109
Sex of Voice x Rater Sex	-0.024	0.081	-0.303	575.009	0.762
Voice Pitch x Profession	-0.001	0.010	-0.074	352.778	0.941
Sex of Voice x Profession	0.010	0.080	0.124	570.816	0.901
Rater Sex x Profession	-0.021	0.080	-0.267	569.225	0.789
Voice Pitch x Sex of Voice x Rater Sex	-0.012	0.021	-0.556	173.570	0.579
Voice Pitch x Sex of Voice x Profession	0.008	0.020	0.397	352.778	0.692
Voice Pitch x Rater Sex x Profession	<0.0001	0.019	0.022	368.741	0.982
Sex of Voice x Rater Sex x Profession	-0.026	0.160	0.161	569.225	0.872
Voice Pitch x Sex of Voice x Rater Sex x Profession	-0.018	0.038	-0.46	368.741	0.644

TABLE 5 Results of our analysis for trustworthiness ratings by speaker voice pitch (z-scored), sex of the speaker (effect coded), rater sex (effect coded), and speaker profession (effect coded), with a confidence interval of 95%.. *p<0.05; **p<=0.01; ***p<=0.0001

FIGURE 6 Scatter plot of our analysis for trustworthiness ratings by speaker voice pitch (z-scored), sex of the speaker, and speaker profession.

DISCUSSION

We tested whether voice pitch affects perceptions of competence, trustworthiness, dominance, and attractiveness among voices of doctors and nurses. We found that voice pitch was negatively associated with dominance. Voice pitch was negatively associated with perceptions of attractiveness and trust in men's voices and positively with perceptions of attractiveness and trust in women's voices. These findings were also qualified by the sex and profession of the vocalizer.

Competence

Although the effect of low voice pitch on perceptions of competence was not significant, voice pitch had a significantly stronger negative effect on perceptions of male doctors' competence than female doctors and nurses or male nurses. Female doctors were judged as more competent doctors than male doctors. Female nurses and male nurses were judged similarly on competence by participants. These results may be representative of participants' experiences in the healthcare field. Potentially representing that women physicians frequently outperform men physicians (C. J. Wallis et al., 2017; C. J. D. Wallis et al., 2023) and that in there day to day experiences patients feel equally cared for my nurses regardless of their sex (Budu et al., 2019).

Attractiveness

Attractiveness judgements were qualified by an interaction between speaker's voice pitch, sex of voice, rater sex and profession. Higher pitched voices were perceived as more attractive than lower pitched voices in female speakers, and lower pitched voices were perceived as more attractive in male speakers. This is consistent with previous work

on voice attractiveness judgements (Borkowska & Pawlowski, 2011; Feinberg et al., 2005, 2008; Hodges-Simeon et al., 2010). The effect of pitch on attractiveness was stronger for male doctors than male nurses and stronger for female nurses than for female doctors. This suggests that for female doctors and male nurses, negative social stereotypes might reduce the effects of voice pitch on attractiveness ratings. We also found that the effects of pitch on attractiveness were stronger for male than female voices, which is also consistent with previous work (see Pisanski & Feinberg, 2018 for review). Furthermore, we also found opposite-sex biases in preferences for voice pitch, where women had stronger preferences for low pitch in men's voices than did men, and men had stronger preferences for high voice pitch in women's voices than did women. These results are also consistent with prior work (Jones et al., 2010).

Male raters judged female speakers as more attractive than male speakers, there was no difference in female judgements of attractiveness. Female speakers were judged to be more attractive than male speakers. Male speakers with higher pitched voices were judged to be less attractive overall.

Dominance

Voice pitch negatively predicted dominance ratings of male doctors and female nurses more than it affected dominance ratings of female doctors and male nurses. This is consistent with the effects of gender stereotypes that we found on competence and dominance ratings. When voices were labelled as having professions counter stereotypical to their genders, the gender mismatch seems to reduce the associations between pitch and dominance.

Lower pitched voices were judged to be more dominant in speakers labelled as doctors and nurses, in both female and male speakers, by both female and male raters. This is consistent with previous results finding that lower pitched voices are generally rated as more dominant sounding. (Anderson & Klofstad, 2012; Armstrong et al., 2019; Hodges-Simeon et al., 2010; McAleer et al., 2014; Puts et al., 2006).

Trustworthiness

We found that men with lower pitched voices and women with higher pitched were judged as relatively more trustworthy. However, these findings were not statistically significant. This is not surprising with how mixed the findings are on the relationship between voice pitch and trustworthiness perceptions. These results could be explained by people perceiving the context of a medical interactions differently. The imagined context of the exchange with a nurse or doctor might influence how participants evaluate the trustworthiness of their voice (O'Connor & Barclay, 2018). If participants understand the exchange with the speaker to be of a professional nature, they might favour low-pitched male voices and high-pitched female voices (McAleer et al., 2014; Montano, Tigue, Isenstein, Barclay, & Feinberg, 2017; O'Connor & Barclay, 2017; Oleszkiewicz et al., 2017; Tigue, Borak, O'Connor, Schandl, & Feinberg, 2012). Alternatively, if they perceived the exchange to be more personal, they might trust high-pitched male and low-pitched female voices more (O'Connor & Feinberg, 2012; O'Connor, Pisanski, Tigue, Fraccaro, & Feinberg, 2014; O'Connor, Re, & Feinberg, 2011). Recognizing the perceived context is necessary to evaluate how trustworthiness judgements are formed from nurses' and doctors' voices.

CONCLUSION

In conclusion, we found that voice pitch affected perceptions of male and female doctors and nurses differently. Stereotypically gender atypical jobs appear to reduce the association between voice pitch and each of the ratings we tested: attractiveness, competence, dominance, and trustworthiness. Identifying these stereotypes is our first step in overcoming them.

REFERENCES

Ambady, N., LaPlante, D., Nguyen, T., Rosenthal, R., Chaumeton, N., & Levinson, W. (2002). Surgeons' tone of voice: A clue to malpractice history. *Surgery*, *132*(1), 5–9. https://doi.org/10.1067/msy.2002.124733

Anderson, R. C., & Klofstad, C. A. (2012). Preference for Leaders with Masculine Voices Holds in the Case of Feminine Leadership Roles. *PLoS ONE*, *7*(12), e51216. https://doi.org/10.1371/journal.pone.0051216

Apicella, C. L., & Feinberg, D. R. (2009). Voice pitch alters mate-choice-relevant perception in hunter–gatherers. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1659), 1077–1082. https://doi.org/10.1098/rspb.2008.1542

Armstrong, M. M., Lee, A. J., & Feinberg, D. R. (2019). A house of cards: Bias in perception of body size mediates the relationship between voice pitch and perceptions of dominance. *Animal Behaviour*, *147*, 43–51. https://doi.org/10.1016/j.anbehav.2018.11.005

Aung, T., & Puts, D. (2020). Voice pitch: A window into the communication of social power. *Current Opinion in Psychology*, *33*, 154–161. https://doi.org/10.1016/j.copsyc.2019.07.028

Baus, C., McAleer, P., Marcoux, K., Belin, P., & Costa, A. (2019). Forming social impressions from voices in native and foreign languages. *Scientific Reports*, *9*(1), 414. https://doi.org/10.1038/s41598-018-36518-6

Belin, P., Boehme, B., & McAleer, P. (2017). The sound of trustworthiness: Acoustic-based modulation of perceived voice personality. *PloS One*, *12*(10), e0185651.

Borkowska, B., & Pawlowski, B. (2011). Female voice frequency in the context of dominance and attractiveness perception. *Animal Behaviour*, *82*(1), 55–59. https://doi.org/10.1016/j.anbehav.2011.03.024

Bradley, G., Sparks, B., & Nesdale, D. (2001). Doctor Communication Style and Patient Outcomes: Gender and Age as Moderators. *Journal of Applied Social Psychology*, *31*(8), 1749–1773. https://doi.org/10.1111/j.1559-1816.2001.tb02749.x

Bruckert, L., Bestelmeyer, P., Latinus, M., Rouger, J., Charest, I., Rousselet, G. A., Kawahara, H., & Belin, P. (2010). Vocal Attractiveness Increases by Averaging. *Current Biology*, *20*(2), 116–120. https://doi.org/10.1016/j.cub.2009.11.034

Budu, H. I., Abalo, E. M., Bam, V. B., Agyemang, D. O., Noi, S., Budu, F. A., & Peprah, P. (2019). "I prefer a male nurse to a female nurse": Patients' preference for, and satisfaction with nursing care provided by male nurses at the Komfo Anokye teaching hospital. *BMC Nursing*, *18*(1), 47. https://doi.org/10.1186/s12912-019-0369-4

Burgoon, M., Birk, T. S., & Hall, J. R. (1991). Compliance and Satisfaction With Physician-Patient Communication An Expectancy Theory Interpretation of Gender Differences. *Human Communication Research*, *18*(2), 177–208. https://doi.org/10.1111/j.1468-2958.1991.tb00543.x

Cartei, V., Oakhill, J., Garnham, A., Banerjee, R., & Reby, D. (2021). Voice Cues Influence Children's Assessment of Adults' Occupational Competence. *Journal of Nonverbal Behavior*, *45*(2), 281–296. https://doi.org/10.1007/s10919-020-00354-y

Collins, S. A. (2000). Men's voices and women's choices. *Animal Behaviour*, *60*(6), 773–780. https://doi.org/10.1006/anbe.2000.1523

Dielissen, P., Bottema, B., Verdonk, P., & Lagro-Janssen, T. (2011). Attention to gender in communication skills assessment instruments in medical education: A review. *Medical Education*, *45*(3), 239–248. https://doi.org/10.1111/j.1365-2923.2010.03876.x

Feinberg, D. R., DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2008). The Role of Femininity and Averageness of Voice Pitch in Aesthetic Judgments of Women's Voices. *Perception*, *37*(4), 615–623. https://doi.org/10.1068/p5514

Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour*, *69*(3), 561–568. https://doi.org/10.1016/j.anbehav.2004.06.012

Hodges-Simeon, C. R., Gaulin, S. J. C., & Puts, D. A. (2010). Different Vocal Parameters Predict Perceptions of Dominance and Attractiveness. *Human Nature*, *21*(4), 406–427. https://doi.org/10.1007/s12110-010-9101-5

Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C., & Vukovic, J. (2008). Integrating cues of social interest and voice pitch in men's preferences for women's voices. *Biology Letters*, *4*(2), 192–194. https://doi.org/10.1098/rsbl.2007.0626

Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C., & Vukovic, J. (2010). A

domain-specific opposite-sex bias in human preferences for manipulated voice pitch. *Animal Behaviour*, *79*(1), 57–62. https://doi.org/10.1016/j.anbehav.2009.10.003

Klofstad, C. A., Anderson, R. C., & Nowicki, S. (2015). Perceptions of Competence, Strength, and Age Influence Voters to Select Leaders with Lower-Pitched Voices. *PLOS ONE*, *10*(8), e0133779. https://doi.org/10.1371/journal.pone.0133779

Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: Voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1738), 2698–2704. https://doi.org/10.1098/rspb.2012.0311

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Long, J. A. (2020). *jtools: Analysis and Presentation of Social Scientific Data* [Computer software]. https://cran.r-project.org/package=jtools

Mast, M. S., Hall, J. A., Klöckner, C., & Choi, E. (2008). Physician Gender Affects How Physician Nonverbal Behavior Is Related to Patient Satisfaction. *Medical Care*, *46*(12), 1212–1218. https://doi.org/10.1097/MLR.0b013e31817e1877

McAleer, P., Todorov, A., & Belin, P. (2014). How Do You Say 'Hello'? Personality Impressions from Brief Novel Voices. *PLoS ONE*, *9*(3), e90779. https://doi.org/10.1371/journal.pone.0090779

Montano, K. J., Tigue, C. C., Isenstein, S. G. E., Barclay, P., & Feinberg, D. R. (2017).

Men's voice pitch influences women's trusting behaviour. Evolution and Human Behavior. http://dx.doi.org/10.101/j.evolhumbehav.2016.10.010

O'Connor, J. J. M., & Barclay, P. (2017). The influence of voice pitch on perceptions of trustworthiness across social contexts. Evolution and Human Behavior, 38, 506–512. https://doi.org/10.1016/j.evolhumbehav.2017.03.001

O'Connor, J. J. M., & Barclay, P. (2018). High voice pitch mitigates the aversiveness of antisocial cues in men's speech. *British Journal of Psychology, 109*, 812–829. https://doi.org/10.1111/bjop.12310.

O'Connor, J. J. M., & Feinberg, D. R. (2012). The influence of facial masculinity and voice pitch on jealousy and perceptions of intrasexual rivalry. Personality and Individual Differences, 52, 369–373. http://dx.doi.org/10.1016/j.paid.2011.10.036.

O'Connor, J. J. M., Pisanski, K., Tigue, C. C., Fraccaro, P. J., & Feinberg, D. R. (2014). Perceptions of infidelity risk predict women's preferences for low male voice pitch in short-term over long-term relationship contexts. *Personality and Individual Differences*, *56*, 73–77. https://doi.org/10.1016/j.paid.2013.08.029

O'Connor, J. J. M., Re, D. E., & Feinberg, D. R. (2011). Voice pitch influences perceptions of sexual infidelity. Evolutionary Psychology, 9, 64–78. http://dx.doi.org/10.1177/ 147470491100900109.

Oleszkiewicz, A., Pisanski, K., Lachowicz-Tabaczek, K., & Sorokowska, A. (2017). Voice-based assessments of trustworthiness, competence, and warmth in blind and sighted adults. *Psychonomic Bulletin & Review*, *24*(3), 856–862. https://doi.org/10.3758/s13423-016-1146-y

Pisanski, K., Feinberg, D., Oleszkiewicz, A., & Sorokowska, A. (2017). Voice cues are used in a similar way by blind and sighted adults when assessing women's body size. *Scientific Reports*, *7*(1), 10329. https://doi.org/10.1038/s41598-017-10470-3

Pisanski, K., & Feinberg, D. R. (2018). *Oxford handbook of voice perception.*

Pisanski, K., Groyecka-Bernard, A., & Sorokowski, P. (2021). Human voice pitch measures are robust across a variety of speech recordings: methodological and theoretical implications. *Biology letters*, *17*(9), 20210356.

Puts, D. A., Gaulin, S. J. C., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior*, *27*(4), 283–296. https://doi.org/10.1016/j.evolhumbehav.2005.11.003

R Core Team. (2021). *R: A Language and Environment for Statistical Computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Re, D. E., O'Connor, J. J. M., Bennett, P. J., & Feinberg, D. R. (2012). Preferences for Very Low and Very High Voice Pitch in Humans. *PLoS ONE*, *7*(3), e32719. https://doi.org/10.1371/journal.pone.0032719

Schattner, A., Rudin, D., & Jellin, N. (2004). Good physicians from the perspective of their patients. *BMC Health Services Research*, *4*(1), 26. https://doi.org/10.1186/1472-6963-4-26

Schild, C., Braunsdorf, E., Steffens, K., Pott, F., & Stern, J. (2022). Gender and Context-

Specific Effects of Vocal Dominance and Trustworthiness on Leadership Decisions. *Adaptive Human Behavior and Physiology*, *8*(4), 538–556. https://doi.org/10.1007/s40750-022-00194-8

Schild, C., Stern, J., & Zettler, I. (2019). Linking men's voice pitch to actual and perceived trustworthiness across domains. *Behavioral Ecology*, arz173. https://doi.org/10.1093/beheco/arz173

Shiramizu, V. K. M., Lee, A. J., Altenburg, D., Feinberg, D. R., & Jones, B. C. (2022). The role of valence, dominance, and pitch in perceptions of artificial intelligence (AI) conversational agents' voices. *Scientific Reports*, *12*(1), 22479. https://doi.org/10.1038/s41598-022-27124-8

Stephens, T., Spevak, R., Rogalin, C. L., & Hirshfield, L. E. (2016). Drawing Doctors vs. Nurses: Gendered Perceptions of Health Professionals. *Journal of the Indiana Academy of the Social Sciences*, *19*(1).

Tigue, C. C., Borak, D. J., O'Connor, J. J. M., Schandl, C., & Feinberg, D. R. (2012). Voice pitch influences voting behavior. Evolution and Human Behavior, 33, 210–216. http://dx.doi. org/10.1016/j.evolhumbehav.2011.09.004.

von Bültzingslöwen, I., Eliasson, G., Sarvimäki, A., Mattsson, B., & Hjortdahl, P. (2006). Patients' views on interpersonal continuity in primary care: A sense of security based on four core foundations. *Family Practice*, *23*(2), 210–219. https://doi.org/10.1093/fampra/cmi103

Vukovic, J., Feinberg, D. R., Jones, B. C., DeBruine, L. M., Welling, L. L. M., Little, A.

C., & Smith, F. G. (2008). Self-rated attractiveness predicts individual differences in women's preferences for masculine men's voices. *Personality and Individual Differences*, *45*(6), 451–456. https://doi.org/10.1016/j.paid.2008.05.013

Vukovic, J., Jones, B. C., Feinberg, D. R., DeBruine, L. M., Smith, F. G., Welling, L. L. M., & Little, A. C. (2011). Variation in perceptions of physical dominance and trustworthiness predicts individual differences in the effect of relationship context on women's preferences for masculine pitch in men's voices. British Journal of Psychology, 102, 37–48. http://dx.doi.org/10.1348/000712610X498750.

Wallis, C. J. D., Jerath, A., Aminoltejari, K., Kaneshwaran, K., Salles, A., Coburn, N., Wright, F. C., Gotlib Conn, L., Klaassen, Z., Luckenbaugh, A. N., Ranganathan, S., Riveros, C., McCartney, C., Armstrong, K., Bass, B., Detsky, A. S., & Satkunasivam, R. (2023). Surgeon Sex and Long-Term Postoperative Outcomes Among Patients Undergoing Common Surgeries. *JAMA Surgery*. https://doi.org/10.1001/jamasurg.2023.3744

Wallis, C. J., Ravi, B., Coburn, N., Nam, R. K., Detsky, A. S., & Satkunasivam, R. (2017). Comparison of postoperative outcomes among patients treated by male and female surgeons: A population based matched cohort study. *BMJ*, j4366. https://doi.org/10.1136/bmj.j4366

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., & Grolemund, G. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Zelek, B., & Phillips, S. P. (2003). Gender and power: Nurses and doctors in Canada.

International Journal for Equity in Health, *2*(1), 1. https://doi.org/10.1186/1475-9276-2-1

Zhang, J., & Reid, S. A. (2017). Aggression in young men high in threat potential increases after hearing low-pitched male voices: Two tests of the retaliation-cost model. *Evolution and Human Behavior*, *38*(4), 513–521. https://doi.org/10.1016/j.evolhumbehav.2017.02.005

Zhang, J., Zheng, L., & Zheng, Y. (2018). Consistency in preferences for masculinity in faces, bodies, voices, and personality characteristics among homosexual men in China. *Personality and Individual Differences*, *134*, 137–142. https://doi.org/10.1016/j.paid.2018.06.009

Zuckerman, M., & Driver, R. E. (1988). What sounds beautiful is good: The vocal attractiveness stereotype. *Journal of Nonverbal Behavior*, *13*(2), 67–82. https://doi.org/10.1007/BF00990791

CHAPTER 5: GENERAL DISCUSSION

This thesis took a multifaceted approach to understanding how we form perceptions of voices. In chapters 2 and 4, we found that voice pitch is a strong predictor of attractiveness and other ratings. We also found in chapter 3 that the ratings we studied in chapters 2 and 4 were important traits that people consciously describe when listening to voices. Although preferences for voices did not produce attractiveness aftereffects in chapter 2, we did find in chapter 4 that altering job title of vocalizers in ways that were inconsistent with gender stereotypes did change voice perception. Voice perceptions are malleable. The plasticity of voice perceptions is likely driven more by social influences than changes in cognitive representation.

The null findings from chapter 2 regarding attractiveness aftereffects, in conjunction with our other work showing that averaging voices decreases distinctiveness ratings but does not change attractiveness ratings (Ostrega et al., 2024) suggests that averageness – at least in the sense suggested by prior work (Bruckert et al., 2010) is not as attractive as previously though. We did however replicate findings that low pitch is attractive in men's voices, and high pitch is attractive in women's voices in chapters 2 and 4 and replicated that finding in our other recent work on averaging voices (Ostrega et al., 2024).

We did not measure voice qualities in chapter 3, but we did ask people to describe voices. People's descriptions of voices mapped tightly on to concepts that people have been studying in voice perception for many years. For example, valence and dominance accounted for about 33% of all topics people described. This is consistent with work on

personality and the voice suggesting that dominance and valence are important (McAleer et al., 2014). While prior work used a questionnaire that seemed to only include questions about dominance and valence (McAleer et al., 2014), our work validates this approach, showing that these are indeed important traits. By looking outside of personality, we also found that people pay attention to people's accents, and other social traits that seem less well studied, but more well-known than other traits.

CHAPTER 2: ADAPTATION TO VOICE ATTRACTIVENESS

In chapter 2 we attempted to extend findings by Little et al. (2013), who showed brief exposure to voices varying in pitch produced normality aftereffects to attractiveness judgements. Prior work in faces has shown that magnitude (Langlois & Roggman, 1990) and direction (DeBruine et al., 2007) from averageness have effects on facial attractiveness. We used identical stimuli and procedures and did not find that adaptation to voices varying in attractiveness judgements produced attractiveness aftereffects, even though we had uneven sample sizes.

Our other study found no evidence that voice attractiveness judgements are tied to voice averageness (Ostrega et al., 2024). Furthermore, prior work on voice pitch and average voices has also never found that average voice pitch is attractive (Feinberg et al., 2008; Re et al., 2012). The one study that found that averaging voices increases their attractiveness (Bruckert et al., 2010) appears to be the anomaly. Indeed, we found preferences for high pitch in women's voices and low pitch in men's voices in chapters 2 and 4. In both studies, there was no indication that average voice pitch was attractive.

LIMITATIONS AND FUTURE DIRECTIONS

Perceptual adaptation of voice pitch judgements need re-examination. Future work should replicate adaptation findings of voice pitch on attractiveness and normality judgments using both male and female voices and raters (Little et al., 2013). Examining these perceptual aftereffects would improve our understanding of the differences and similarities between attractiveness and normality impressions from the voice.

This study did not examine the duration of perceptual aftereffect for voice pitch on attractiveness judgements. Manipulating the length of time participants provide judgements after adaptation to both voice attractiveness and normality judgements would improve measurements of the presence and duration of any perceptual aftereffects (Zäske et al., 2009, 2010). This method could also improve our understand of how different traits might produce aftereffects of varying lengths based on where the trait resides in a temporo-perceptual hierarchy (Lavan, 2023b).

CHAPTER 3: UNCONSTRAINED FIRST IMPRESSIONS OF VOICES

In chapter 3 we replicated and extended research by Lavan (2023a) who used free form responses to represent person perceptions from faces and voices to represent the impressions and judgments people obtain from faces and voices. Here, we focused on voices, and instead of only using six stimuli, we used 100. The other key difference in our studies is that we used Latent Dirichlet Allocation to analyze the key topics that were described by raters. Our results were consistent with Lavan (2023a). In both chapter 3 and Lavan (2023a), participants described personality and physical traits the most, followed by social characteristics and finally voice quality. One third of our descriptions were

summarized using valence and dominance attributes which highlighted the importance of these dimensions in person judgements as previously stated by Oosterhoff & Todorov (2008). Like Lavan (2023a) we also found that voice impressions are not limited to personality characteristics and the free descriptions provide a fuller account of voice impressions.

LIMITATIONS AND FUTURE DIRECTIONS

This study was limited by the small number of voice stimuli and exclusive recruitment of an undergraduate student sample providing relatively short descriptions. To address this, we have recently finished collecting data on a follow up study using a large sample (n>1000) of voices form the VoxCeleb database, that contains “speech from the wild” or ambient voice recordings from YouTube videos of celebrity interviews. We found very similar topics despite the differences in sampling.

The next step in exploring the perceptual voice space from these findings is to take the key descriptors that we found and have those validated by an independent sample of raters (Oosterhof & Todorov, 2008). Once that is complete the ratings would then need to be run through an oblique factor rotation to evaluate how they summarize in dimensional space (McAleer et al., 2014; Oosterhof & Todorov, 2008; Sutherland et al., 2013).

Once, those steps are complete, an acoustical analysis of these voices would be used to evaluate if any components or characteristics of the voices correlate with our found summarizing dimensions. If for example some voices characteristics consistently sounds more trustworthy than others if trustworthiness were to be one of the summarizing

dimensions. If this is the case, we would need to evaluate if changing voices on those same acoustical components can alter perceptions of individuals from their voice.

CHAPTER 4: THE EFFECTS OF VOICE PITCH ON JUDGMENTS OF ATTRACTIVENESS, COMPETENCE, DOMINANCE, AND TRUSTWORTHINESS OF DOCTORS AND NURSES

After finding the importance of valence and dominance attributes in our free descriptions of voices, we wanted to see if trust, which is highly correlated with valence and dominance, apply to medical scenarios since attractiveness, competence, dominance, and trustworthiness are all important factors in delivering patient care. People perceived low pitched men's and women's voices as more dominant than higher pitched voices (Anderson & Klofstad, 2012; Klofstad et al., 2012; Tigue et al., 2012), we observed effects in the same directions for competence ratings, but these effects were not significant. Lower pitched men's voices and higher pitched women's voices were perceived as more attractive (Apicella & Feinberg, 2009; Borkowska & Pawlowski, 2011; Cussigh et al., 2020; Feinberg et al., 2005, 2008, 2011; Jones et al., 2010; Re et al., 2012; Schild et al., 2020; Simmons et al., 2011) and trustworthy (Belin et al., 2017; McAleer et al., 2014; Oleszkiewicz et al., 2017; Schild et al., 2019) than were higher pitched men's voices and lower pitched women's voices. We replicated the findings that low-pitched voices were rated more dominant than high-pitched voices this was true across all professions, sex of voice and raters. We also replicated the finding that high-pitched female voices and low-pitched male voices are judged as most attractive (Apicella & Feinberg, 2009; Borkowska & Pawlowski, 2011; Cussigh et al., 2020; Feinberg et al.,

2005, 2008, 2011; Jones et al., 2010; Re et al., 2012; Schild et al., 2020; Simmons et al., 2011). Although men with lower pitched voices and women with higher pitched voices were judged as slightly more trustworthy, these results were not statistically significant.

These findings highlight how judgements made by others can influence how they perceive someone else, based on no other information than a label and the sound of their voice. We found that gender stereotyping interacts with voice perception in a healthcare context. Our results show how patients might form judgements of members of their healthcare team. Future work could look at how this might influence how care is received by a patient.

One potential limitation was using fake doctor and nurse voices. It is possible that our stimuli were not convincing, affecting results. This is unlikely given that the pattern of results we found is consistent with prior work (Armstrong et al., 2019; Belin et al., 2017; Borkowska & Pawlowski, 2011; Feinberg et al., 2005, 2008; Klofstad et al., 2015; McAleer et al., 2014; Oleszkiewicz et al., 2017; Puts et al., 2006; Re et al., 2012; Schild et al., 2019, 2022; Simmons et al., 2011).

FINAL CONCLUSIONS

This dissertation explored the rich set of information that people retrieve from listening to the sound of voices. We found support for the idea that voice pitch is a key acoustic feature in social perception of voices, whereas averageness played little role in attractiveness ratings. People's unconstrained ratings of voices were consistent with studies using questionnaires such as (McAleer et al., 2014; Oosterhof & Todorov, 2008; Sutherland et al., 2013) and chapters 2 and 4.

Although a lot of these findings are starting points to exploring the complexities of the social voice space and how judgments in the space might affect human interactions they work together to demonstrate the richness of information that can be collected from the voice. Voices might not exactly be the auditory face that others have claimed (Belin et al., 2011), but they are still of incredible importance in human interactions especially when it comes to the formation of impressions of others.

REFERENCES

Anderson, R. C., & Klofstad, C. A. (2012). Preference for Leaders with Masculine Voices Holds in the Case of Feminine Leadership Roles. *PLoS ONE*, *7*(12), e51216. https://doi.org/10.1371/journal.pone.0051216

Apicella, C. L., & Feinberg, D. R. (2009). Voice pitch alters mate-choice-relevant perception in hunter–gatherers. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1659), 1077–1082. https://doi.org/10.1098/rspb.2008.1542

Armstrong, M. M., Lee, A. J., & Feinberg, D. R. (2019). A house of cards: Bias in perception of body size mediates the relationship between voice pitch and perceptions of dominance. *Animal Behaviour*, *147*, 43–51. https://doi.org/10.1016/j.anbehav.2018.11.005

Belin, P., Bestelmeyer, P. E. G., Latinus, M., & Watson, R. (2011). Understanding Voice Perception: Understanding voice perception. *British Journal of Psychology*, *102*(4), 711–725. https://doi.org/10.1111/j.2044-8295.2011.02041.x

Belin, P., Boehme, B., & McAleer, P. (2017). The sound of trustworthiness: Acoustic-based modulation of perceived voice personality. *PloS One*, *12*(10), e0185651.

Borkowska, B., & Pawlowski, B. (2011). Female voice frequency in the context of dominance and attractiveness perception. *Animal Behaviour*, *82*(1), 55–59. https://doi.org/10.1016/j.anbehav.2011.03.024

Bruckert, L., Bestelmeyer, P., Latinus, M., Rouger, J., Charest, I., Rousselet, G. A., Kawahara, H., & Belin, P. (2010). Vocal Attractiveness Increases by Averaging. *Current Biology*, *20*(2), 116–120. https://doi.org/10.1016/j.cub.2009.11.034

Cussigh, G., Ballester-Arnal, R., Gil-Llario, M. D., Giménez-García, C., & Castro-Calvo, J. (2020). Fundamental frequency of the female's voice: A cross-country empirical study on its influence on social and sexual selection. *Personality and Individual Differences*, *160*, 109937. https://doi.org/10.1016/j.paid.2020.109937

DeBruine, L. M., Jones, B. C., Unger, L., Little, A. C., & Feinberg, D. R. (2007). Dissociating averageness and attractiveness: Attractive faces are not always average. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(6), 1420–1430. https://doi.org/10.1037/0096-1523.33.6.1420

Feinberg, D. R., DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2008). The Role of Femininity and Averageness of Voice Pitch in Aesthetic Judgments of Women's Voices. *Perception*, *37*(4), 615–623. https://doi.org/10.1068/p5514

Feinberg, D. R., Jones, B. C., DeBruine, L. M., O'Connor, J. J. M., Tigue, C. C., & Borak, D. J. (2011). Integrating fundamental and formant frequencies in women's preferences for men's voices. *Behavioral Ecology*, *22*(6), 1320–1325. https://doi.org/10.1093/beheco/arr134

Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour*, *69*(3), 561–568. https://doi.org/10.1016/j.anbehav.2004.06.012

Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C., & Vukovic, J. (2010). A

domain-specific opposite-sex bias in human preferences for manipulated voice pitch. *Animal Behaviour*, *79*(1), 57–62. https://doi.org/10.1016/j.anbehav.2009.10.003

Klofstad, C. A., Anderson, R. C., & Nowicki, S. (2015). Perceptions of Competence, Strength, and Age Influence Voters to Select Leaders with Lower-Pitched Voices. *PLOS ONE*, *10*(8), e0133779. https://doi.org/10.1371/journal.pone.0133779

Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: Voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1738), 2698–2704. https://doi.org/10.1098/rspb.2012.0311

Lavan, N. (2023a). How do we describe other people from voices and faces?. *Cognition*, *230*, 105253.

Lavan, N. (2023b). The time course of person perception from voices: A behavioral study. *Psychological Science*, *34*(7), 771-783.

Langlois, J. H., & Roggman, L. A. (1990). Attractive Faces Are Only Average. *Psychological Science*, *1*(2), 115–121. https://doi.org/10.1111/j.1467-9280.1990.tb00079.x

McAleer, P., Todorov, A., & Belin, P. (2014). How Do You Say 'Hello'? Personality Impressions from Brief Novel Voices. *PLoS ONE*, *9*(3), e90779. https://doi.org/10.1371/journal.pone.0090779

Oleszkiewicz, A., Pisanski, K., Lachowicz-Tabaczek, K., & Sorokowska, A. (2017).

Voice-based assessments of trustworthiness, competence, and warmth in blind and sighted adults. *Psychonomic Bulletin & Review*, *24*(3), 856–862. https://doi.org/10.3758/s13423-016-1146-y

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Ostrega, J., Shiramizu, V., Lee, A. J., Jones, B. C., & Feinberg, D. R. (2024). No evidence that averaging voices influences attractiveness. *Scientific Reports*, *14*(1), 10488.

Puts, D. A., Gaulin, S. J. C., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior*, *27*(4), 283–296. https://doi.org/10.1016/j.evolhumbehav.2005.11.003

Re, D. E., O'Connor, J. J. M., Bennett, P. J., & Feinberg, D. R. (2012). Preferences for Very Low and Very High Voice Pitch in Humans. *PLoS ONE*, *7*(3), e32719. https://doi.org/10.1371/journal.pone.0032719

Schild, C., Aung, T., Kordsmeyer, T. L., Cardenas, R. A., Puts, D. A., & Penke, L. (2020). Linking human male vocal parameters to perceptions, body morphology, strength and hormonal profiles in contexts of sexual selection. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-77940-z

Schild, C., Braunsdorf, E., Steffens, K., Pott, F., & Stern, J. (2022). Gender and Context-Specific Effects of Vocal Dominance and Trustworthiness on Leadership Decisions. *Adaptive Human Behavior and Physiology*, *8*(4), 538–556. https://doi.org/10.1007/s40750-022-00194-8

Schild, C., Stern, J., & Zettler, I. (2019). Linking men's voice pitch to actual and perceived trustworthiness across domains. *Behavioral Ecology*, arz173. https://doi.org/10.1093/beheco/arz173

Simmons, L. W., Peters, M., & Rhodes, G. (2011). Low Pitched Voices Are Perceived as Masculine and Attractive but Do They Predict Semen Quality in Men? *PLoS ONE*, *6*(12), e29271. https://doi.org/10.1371/journal.pone.0029271

Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Michael Burt, D., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition*, *127*(1), 105–118. https://doi.org/10.1016/j.cognition.2012.12.001

Tigue, C. C., Borak, D. J., O'Connor, J. J. M., Schandl, C., & Feinberg, D. R. (2012). Voice pitch influences voting behavior. *Evolution and Human Behavior*, *33*(3), 210–216. https://doi.org/10.1016/j.evolhumbehav.2011.09.004

Zäske, R., Schweinberger, S. R., Kaufmann, J. M., & Kawahara, H. (2009). In the ear of the beholder: Neural correlates of adaptation to voice gender. *European Journal of Neuroscience*, *30*(3), 527–534. https://doi.org/10.1111/j.1460-9568.2009.06839.x

Zäske, R., Schweinberger, S. R., & Kawahara, H. (2010). Voice aftereffects of adaptation to speaker identity. *Hearing Research*, *268*(1–2), 38–45. https://doi.org/10.1016/j.heares.2010.04.011