2023

Does implicit racial bias affect auditory-perceptual evaluations of dysphonic voices?

BOSTON UNIVERSITY

SARGENT COLLEGE OF HEALTH AND REHABILITATION SCIENCES

Thesis

DOES IMPLICIT RACIAL BIAS AFFECT AUDITORY-PERCEPTUAL EVALUATIONS OF DYSPHONIC VOICES?

by

RACHEL LOWELL NOROTSKY

B.A., Wake Forest University, 2017

Submitted in partial fulfillment of the

requirements for the degree of

Master of Science

2023

Approved by

First Reader

Cara E. Stepp, Ph.D.
Professor of Speech, Language & Hearing Sciences
Sargent College of Health and Rehabilitation Sciences

Professor of Biomedical Engineering
Boston University, College of Engineering

Professor of Otolaryngology - Head and Neck Surgery
Boston University, School of Medicine

Second Reader

Cara Sauder, Ph.D., CCC-SLP
Lecturer of Speech and Hearing Sciences
University of Washington

Third Reader

Kimberly Dahl, M.S. CCC-SLP
Doctoral Research Assistant
Stepp Lab for Sensorimotor Rehabilitation Engineering

Fourth Reader

Lauren F. Tracy, MD
Assistant Professor of Otolaryngology - Head & Neck Surgery
Boston University, Chobanian & Avedisian School of Medicine

ACKNOWLEDGMENTS

I would first like to thank Dr. Cara Stepp for welcoming me into her lab before I began the MS-SLP program and for providing such valuable thesis supervision and career guidance ever since. I am so grateful that our research interests aligned with this topic. I would also like to thank Kimberly Dahl for her close mentorship and expertise throughout this project. You always encouraged me to strive for excellence and I have learned so much from you. This project would not have been possible without such phenomenal mentors.

I would also like to extend my gratitude to Dr. Cara Sauder and Dr. Lauren Tracy for participating on my thesis committee. Dr. Sauder brought incredible knowledge of perceptual bias in voice evaluations to this project based on her previous publications. Dr. Tracy's perspective as a laryngologist was invaluable in our discussions of study results and their clinical impact. I appreciate the wealth of experience you both brought to this project as current experts in the field.

I would like to thank all members of the Stepp Lab, especially Sarah Cocroft and Megan Cushman who both played integral roles in the project's data collection. Sarah devoted her time to being trained in the study's protocol so she could enroll participants and supervise study sessions. Megan went to great lengths to schedule all participants for the experiment and played an essential role as the Stepp Lab manager. I would also like to thank Daniel Buckley and Dr. Katherine Marks for being the professors to spark my enthusiasm for all things voice during my first semester of graduate school. You have served as my role models ever since and I feel immensely grateful to have learned from

both of you.

I am also incredibly grateful for my cohort, the Boston University MS-SLP Class of 2023, for the camaraderie that kept me motivated over the last two years. My hope for this project was for it to be clinically relevant, and I am so appreciative of my fantastic clinical supervisors for helping me grow into a confident and skilled clinician. Thank you to all of my professors in the SLHS department for cheering me on throughout this project and for reminding me why I love this field. Lastly, I would not be at this stage of my career without my parents, Mitchell and Deborah, and my sisters, Allison and Elise. Thank you for your unconditional love throughout graduate school and for being the greatest supporters in all of my endeavors.

DOES IMPLICIT RACIAL BIAS AFFECT AUDITORY-PERCEPTUAL EVALUATIONS OF DYSPHONIC VOICES?

RACHEL LOWELL NOROTSKY

ABSTRACT

Purpose: The purpose of this study was to further understand the role of implicit racial bias in auditory-perceptual evaluations of dysphonic voices. This was assessed by determining if a biasing effect exists for novice raters in their auditory-perceptual ratings of Black and White speakers of General American English who have voice disorders.

Method: 30 novice clinicians at Boston University listened to audio files of 20 Black speakers and 20 White speakers of General American English with voice disorders. Across two study sessions, listeners rated the overall severity of each voice heard using the CAPE-V visual analog scale and completed the Harvard Implicit Association Test (IAT).

Results: Both Black and White speakers were rated as less severe on the CAPE-V when their race was labeled as Black. No significant relationship was found between Harvard IAT scores and severity ratings.

Conclusion: These findings suggest a minimizing bias for Black patients with voice disorders who present to speech-language pathologists for evaluation. These results will contribute to the understanding of how demographic information unrelated to a voice disorder may impact a patient's visit with a clinician.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

The annual prevalence of voice disorders among adults in the US is 6–7% (Bainbridge et al., 2017; Bhattacharyya, 2014). A voice disorder is defined by the American Speech-Language-Hearing Association (ASHA) as "when an individual expresses concern about having a voice that does not meet their daily needs given differences in vocal quality, pitch, loudness, resonance and/or duration" (ASHA, 1993). They may result from physical changes to the vocal mechanism (e.g., vocal fold nodules), inefficient use of the vocal mechanism despite normal structures (e.g., muscle tension dysphonia), neurologic causes, or psychogenic causes such as anxiety or chronic stress (Ferrand, 2012). Voice disorders are seen across the lifespan and are frequently evaluated by speech-language pathologists using auditory-perceptual measures, one of several components of a voice evaluation.

Auditory-perceptual evaluations of voice are critical within the voice disorders specialty of speech-language pathology. The Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) is an instrument commonly used for this by clinicians to standardize evaluations (Kempster et al., 2009). The CAPE-V assesses six features of dysphonia: overall severity, roughness, breathiness, strain, pitch, and loudness. Research has shown the CAPE-V to be both valid and reliable for those trained in voice disorders (Helou et al., 2010; Zraick et al., 2011).

However, auditory-perceptual methods of voice evaluation are subject to biases. For example, perceptual biases in voice evaluations arise when information beyond just acoustic data is presented. There is evidence of medical diagnosis knowledge biasing

auditory-perceptual voice evaluations. In a study examining whether diagnostic information impacted auditory-perceptual evaluations of voice, researchers found that listeners consistently rated voices as more severe on the measures of roughness and breathiness in readings of The Rainbow Passage when they knew the speaker's medical diagnosis (Eadie et al., 2011). This suggests that knowledge of medical diagnosis may bias clinicians in their auditory-perceptual evaluations of dysphonic voices (Eadie et al., 2011). Sauder and Eadie (2020) found that including a diagnosis impacted how raters assessed dysphonia in readings of the Rainbow Passage, whether the diagnosis was accurate or not. They found that inaccurate diagnoses showed a larger main effect than accurate diagnoses. This study found evidence for both an overpathologizing bias as well as a minimizing bias that was contingent upon the feature being assessed and the accuracy of the diagnosis (Sauder & Eadie, 2020). It showed that, overall, speech-language pathology graduate students are susceptible to bias in voice evaluations. This evidence suggests that perceptual bias exists in auditory-perceptual evaluations of voice when patient information is given, though there is little research on racial bias specifically.

Implicit racial biases are subconscious and often happen automatically, in contrast to overt racism. They are "mental associations made between members of a social group who share a particular feature and certain attributes or evaluations" (FitzGerald et al., 2019, p. 1). Implicit biases are difficult to address with targeted awareness as they reside in our subconscious, therefore we do not know that they are occurring (Dalton & Villagran, 2018). In fact, they often surprise us by contradicting our declared beliefs.

In order to address how implicit racial bias may influence voice evaluations, it must first be known whether race creates inherent differences in voice. Race is a social construct, therefore it should have no relationship to biological differences in the vocal mechanism. In a study that paired vowel productions of Black and White speakers (adult males and females) based on age, height, and weight, no significant differences were found between voices on aerodynamic nor acoustic measures, except for maximum-declination rate such that White speakers had higher rates of declination (Sapienza, 1997). This suggests that Black speakers and White speakers when paired on demographics should sound comparable during vowel production in terms of their acoustic characteristics. This finding was supported in another study that compared Black and White elderly speakers' vocal productions and found no differences between race on the acoustic features of voice (Xue & Fucci, 2000). This provides rationale that no acoustic characteristics of voice should imply a speaker's race. There is no current evidence of physiological nor acoustic differences in the voices of speakers that is contingent upon their race when paired on other demographics. Given that there is no biological basis for race-based differences in voice, any difference in severity ratings of White and Black speakers when paired on other factors may likely be the result of the rater's implicit racial biases.

Studies of implicit bias in healthcare have spotlighted that patients of color are evaluated, diagnosed, and ultimately treated differently from White patients. In 2003, Smedley et al. published a book titled *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care,* which inspired a variety of healthcare fields to reflect

on how provider biases had contributed to these prominent healthcare disparities. A systematic review examining implicit bias in healthcare professionals found that 35 out of 42 eligible articles reported notable levels of implicit bias in healthcare professionals, and that the majority of the 42 articles examined racial/ethnic biases (FitzGerald & Hurst, 2017). All eligible articles that compared implicit bias to quality of care found a significant negative relationship between the two (FitzGerald & Hurst, 2017).

Another systematic review focused specifically on implicit racial/ethnic bias in healthcare providers used an implicit association task to find that provider bias consistently correlated with worse interactions between patients and providers (Maina et al., 2018). One study found implicit racial bias to impact pain management trajectories such that pediatricians showing increased pro-White implicit bias also showed decreased narcotic medication prescription for Black patients (Sabin & Greenwald, 2012). In a study in which physicians reported no explicit racial bias, implicit measures still found a significant preference for White Americans amongst these physicians, along with implicit stereotypes that Black Americans are less cooperative during medical visits and in general (Green et al., 2007). In this same study, pro-White implicit bias correlated with increased likelihood of not treating Black patients with thrombolysis, yet treating White patients for this same condition (Green et al., 2007).

Racial disparities in medical evaluations have been well-documented for decades, with examples such as Black patients being less likely to receive commonly prescribed medications and angiograms when in the hospital after a heart attack (Arora et al., 2018). Rathore et al. (2000) found that medical students showed racial bias in their evaluation of

a Black female versus a White male both presenting with identical angina symptoms such that the Black female was viewed to have a "less desirable health state" than the White male. These students were also less likely to diagnose the Black woman with angina, even though both individuals presented with angina. This suggests that although both patients presented the same, the White male's symptoms were taken more seriously, and the Black female's presentation was viewed as less severe (Rathore et al., 2000). Studies like this suggest that nonmedical factors such as implicit racial biases may impact diagnostic evaluations.

Research focused on implicit racial biases is needed given the discrepancies in care of White and Black patients. A review of literature focused on implicit biases in pediatric healthcare found clinician bias to negatively impact Black youth specifically (Mulchan et al., 2021). Another review identified that Black patients are less likely to receive care that adheres to clinical guidelines, even after accounting for socioeconomic factors (Odonkor et al., 2021). An example of this is that Black patients were less likely to receive invasive cardiac procedures (Odonkor et al., 2021). Given the well-documented likelihood of Black patients' conditions to not be viewed as severely or taken as seriously as White patients (Javitt et al., 1991; Johnson-Jennings et al., 2020; Odonkor et al., 2021; Rathore et al., 2000), research is needed to see if these findings transfer to the field of speech-language pathology. If so, it would not be surprising to find Black patients with voice disorders to be rated as less severe than White patients, even with the same presentation.

Within the field of speech-language pathology, clinician racial bias has been

broadly investigated, though there is a lack of research assessing its impact within the specialty field of voice disorders. In a study where trained speech-language pathologists and untrained listeners rated the accuracy of children's productions of words with distinct pronunciations in African American English (AAE), productions were rated more accurately when participants were primed with a Black child's face compared to a White child's face (Evans et al., 2018). One study reported that Black students were less likely to receive services for specific language impairments when examining disparities in special education (Morgan et al., 2017). Another study revealed a theme of Black clients with aphasia reporting dissatisfaction over only being recommended for a limited frequency of speech therapy services even though they had large communicative impairments (Mahendra & Spicer, 2014). This study also reported that 80% of the Black clients with aphasia surveyed had at least one negative treatment experience with a speech-language pathologist (Mahendra & Spicer, 2014).

There is an intersection of racial bias and linguistic bias within speech-language pathology. In one study, an English speaker's speech was rated as less standard by undergraduate students when it was paired with an Asian woman's face than when it was paired with a White woman's face, though both photographs were entirely similar except for subject race (Rubin, 1992). This image priming impacted students' perception of accent differences that did not actually exist (Rubin, 1992). The intersection of racism and dialectal discrimination in speech-language pathology has been harmful throughout the years, such that students of color have been overrepresented in special education if their communication patterns do not align with the mainstream language used in United

States public schools (Farrugia-Bernard, 2018). One study investigating the influence of linguistic bias on speech-language pathologists' clinical decision-making found that speech-language pathologists with negative attitudes towards diversity and dialectal variation were more likely to diagnose a client with a communicative disorder over a dialectal difference (Easton & Verdon, 2021). Speech-language pathologists' linguistic bias towards speakers of nonstandard dialects impacts their ability to make objective clinical decisions on whether a client's speech is different or disordered. There are now specific ways to target dialectal bias in the field of speech-language pathology, such as resources outlining which aspects of dialect variation are dialectal differences versus disordered speech (Kester, 2014). Nevertheless, in a field where evaluations of language are intended to be objective for diagnostics, there is clear evidence of rater bias interrupting this through the way clients are perceived beyond just their language.

There is no outline for how to eradicate implicit racial bias in the field of speech-language pathology, though implicit biases have been shown to have a large impact on clinical decision-making. In a commentary published from the *American Journal of Speech-Language Pathology*, authors stated that implicit bias permeates through the field by influencing perception, memory, and behavior (Ellis et al., 2021). A common strategy put forward to improve issues of racial bias is for individuals to reflect on and attempt to minimize their own race-related biases, hence why many graduate programs now require implicit bias training. Although speech-language pathology training programs now recognize the impact of implicit biases and are now attempting to provide implicit bias training as of recent years (Bureau of Professional Licensing, 2021), research is still

needed to understand the extent to which implicit biases impact clients.

A current focus within the field of speech-language pathology is providing antiracist and equitable services for all clients served. The field lacks racial diversity with 91.7% of ASHA members being White (ASHA, 2019). This indicates a mismatch from the racial breakdown of the United States population (*U.S. Census Bureau QuickFacts*, 2021) and is likely a source of considerable bias in the field. In ASHA's *Response to Racism* position statement (ASHA, 2020), the organization stated that by 2025 it aims to "have inclusive policies and practices in place through initiatives such as increasing the diversity of membership and increasing the cultural competence of members." Although it is important to provide overt initiatives for eliminating racism in the field, this does not address the highly relevant issue of implicit bias in speech-language pathology. ASHA's *Issues in Ethics* statement (ASHA, 2017) requires that clinicians approach all client relationships with "awareness, knowledge, and skills about their own culture and cultural biases, strengths, and limitations". This requires an understanding of one's own implicit racial biases and how they impact one's clinical practice.

Purpose of the Current Study

The purpose of the current study was to further understand the role of implicit racial bias in auditory-perceptual evaluations of dysphonic voices. This was assessed by determining if a biasing effect exists for novice raters in their auditory-perceptual ratings of Black and White speakers of General American English who have voice disorders. The focus of the study is on the race-related implicit biases that may arise based on race labels accompanying voice samples. Novice participants were asked to rate the overall severity

of voice samples using the CAPE-V with their only knowledge of each voice being sex, race, and age of the speaker. Race was labeled accurately 50% of the time and inaccurately for the remaining 50% to determine how race labeling subconsciously influenced severity ratings. We anticipated a main effect of race labeling, regardless of underlying speaker race. We hypothesized that speakers would be rated as less severely dysphonic by novice raters on the CAPE-V when their race was labeled as Black. This study also tested for an interaction effect to see if this was true only for speakers of one underlying race, though we did not anticipate a significant interaction given our hypothesis that race labeling would influence ratings regardless of true speaker race. Additionally, we hypothesized that the difference in mean ratings according to race labeling (Labeled Black - Labeled White) would negatively correlate with Harvard Implicit Association Test (IAT) scores. There is no biological reason why race should impact severity ratings, therefore implications of these findings may provide insight into the role of implicit racial bias in clinical settings.

Methods

Voice Recordings

Audio files were selected from the STEPP lab's previous study database. Audio files consisted of two sustained vowels (/a/ and /i/) and the following two sentences from *The Rainbow Passage*: "The rainbow is a division of white light into many beautiful colors. These take the shape of a long round arch, with its path high above and its two ends apparently beyond the horizon" (Fairbanks, 1960). These files were concatenated to

make a single speech sample per speaker that was normalized for peak intensity using a MATLAB script (version 9.10.0.1613233 [R2021a]; The MathWorks, Inc., Natick, MA).

Speakers

The current study included only speakers of General American English in order to eliminate dialect or accented speech as a confound. Language classifications were made based on samples of spontaneous participant speech. African American English was identified using specific linguistic characteristics outlined in research (Charity, 2008). After one listener involved in the study rated the language classification of all eligible speakers, a second listener involved in the study performed the same task, blinded to the ratings of the first listener. Where there were discrepancies between both raters, a third rater involved in the study made the final decision. Once a list of Black speakers of General American English was formed, each speaker was paired to a White speaker in the STEPP lab database. Speakers all had been diagnosed with a voice disorder by a laryngologist and were matched on age, sex, and overall severity ratings that were completed by a voice-specialized speech-language pathologist using the CAPE-V in order to minimize differences between voices. This was intended to allow for interpretation of differences in study ratings to be an effect of race labeling instead of alternative factors. All White speakers were within 5 years of age of a paired Black speaker. There were 20 Black speakers (17 females, 3 males, M = 40.1 years, SD = 17.2 years, range = 18 - 67 years) and 20 White speakers (17 females, 3 males, M = 39.6 years, SD = 16.7 years, range = 19 - 66 years). The average overall severity of dysphonia of Black speakers was 20.9 out of 100 (SD = 12.9, range = 4.8 - 50.7), and the average

severity of White speakers was 20.0 out of 100 (*SD* = 17.4, range = 3.4 - 82.0).

Listeners

30 individuals participated as listeners in the current study. Listeners had all taken a graduate level voice disorders class at Boston University within 14 months of the study. All listeners were in the process of obtaining their Master's degree in speech-language pathology (MS-SLP) from Boston University. The study listeners were all speakers of English who presented without neurological impairment and without speech, voice, language, or hearing disorders. Listeners were financially compensated for their time. They were presented with the opportunity to participate in a paid remote research study for the STEPP lab but were otherwise blinded to the study purpose. Informed consent was obtained from all participants prior to the study, in compliance with the Boston University Institutional Review Board.

Experimental Conditions

Session 1 Ratings. In the first study session, listeners were randomly assigned to one of two conditions. 15 listeners were provided with entirely accurate race labels for each voice heard and 15 listeners were provided with entirely inaccurate race labels for each voice heard. Each listener heard all 40 speakers in Session 1. 20% of audio samples were repeated to measure listener reliability. Given that there were 40 speakers, 8 speaker samples were heard twice in each session, resulting in each listener completing 48 ratings per session. The order of accurate-inaccurate sessions was counterbalanced across listeners such that 50% of listeners received 48 voices with accurate race labels in their first session.

Session 2 Ratings. Four to seven days after their first study session, both groups of listeners returned to complete the same evaluations but with the opposite accuracy condition of their first session. Each listener once again heard all 40 speakers in Session 2 with 8 repeated for reliability, resulting in each listener completing another 48 ratings. Participants who first received 48 accurate race labels now received 48 inaccurate race labels and vice versa.

Listening Procedure

The listening procedure was completed remotely by all participants using the online behavioral research platform Gorilla Experiment Builder (gorilla.sc). Listeners were instructed to complete both study sessions in a quiet environment on their device of choice using their own headphones for improved audition. Listening sessions were supervised remotely by a STEPP lab member in order to solve technological problems and answer listener questions. Each listening session began with a volume adjustment and headphone screening task from the Gorilla open materials repository (Milne et al., 2021). The volume adjustment allowed listeners to adjust their device volume to a comfortable level during the presentation of a 4-second clip of white noise. The headphone screening consisted of a Huggin's pitch task in which listeners identified which of three segments of white noise contained a tone. This confirmed that all listeners used headphones for the listening procedure (Milne et al., 2021).

Once listeners passed the volume adjustment and the headphone screening, they began the primary listening task in which they were presented with the voice recordings outlined previously. Participants rated 48 voices during each listening session: the 20

unique audio files of Black speakers, the 20 unique audio files of White speakers, and 8 of the 40 audio files that were repeated at the end of the listening task to measure intrarater reliability. Each voice recording was presented with the speaker's sex, age, and their labeled race (see Figure 1). Listeners rated the overall severity of each speaker's voice using an online version of the CAPE-V's visual analog scale (VAS) of voice deviance (severity) ranging from mildly deviant, to moderately deviant, to severely deviant. The VAS scale ranged from 0-100 with an anchor at 10 for mild, 35 for moderate, and 72 for severe to align with the anchor positions on the current printable version of the CAPE-V form available from ASHA (asha.org/form/cape-v/). The CAPE-V was modified for online administration and constructed in Gorilla. Once each listener had entered their overall severity rating for one speaker, they were presented with the next speaker's audio file until they had rated all 48 samples. Audio files were presented in a randomized order. Listeners were instructed to listen to the entire audio file at least once before rating the speaker's severity. Listeners were able to play each audio file for a second time in case they experienced a distraction in their environment during initial presentation of the recording. Each listening session took approximately one hour.

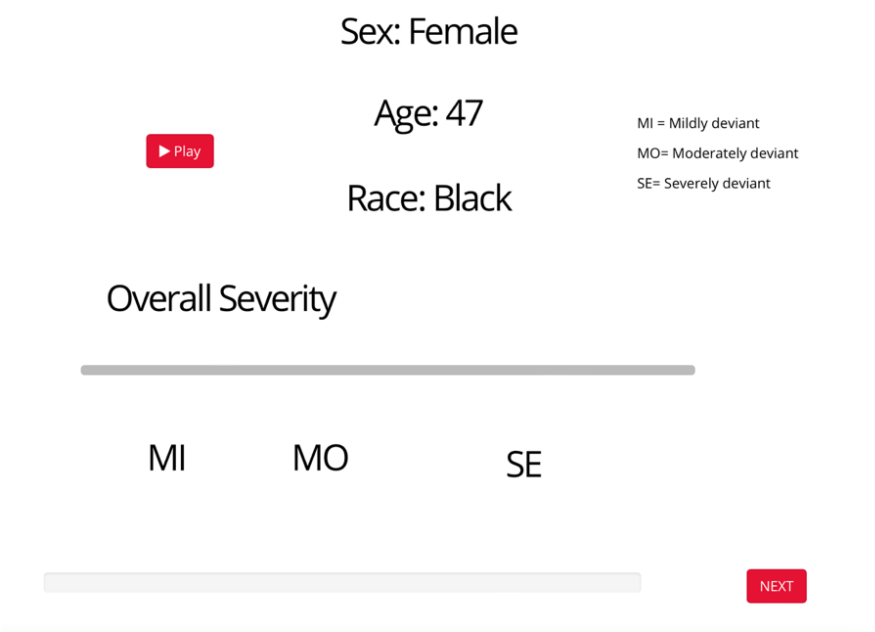

Figure 1. Gorilla experiment display for voice recordings, visual analog scale, and demographic information about voice recordings.

Reliability

Reliability on CAPE-V ratings was completed using intrarater values to measure individual rater reliability on the same trials repeated and interrater values to measure reliability across all raters. Intrarater reliability was calculated using an intraclass correlation coefficient (ICC). Interrater reliability was calculated using a Pearson correlation (*r*). Both of these were interpreted such that values less than 0.5 indicated poor reliability, values between 0.50 and 0.75 indicated moderate reliability, values between 0.75 and 0.90 indicated good reliability, and values greater than 0.90 indicated excellent reliability. (Koo & Li, 2016)

Implicit Bias Task

As the final task of the second session's experimental procedure, participants completed the Harvard Implicit Association Test (IAT) in Gorilla as a measure of their implicit racial bias (Greenwald et al., 1998). This is a commonly used pre-existing measure that asks participants to sort pictures and words into groups as fast as they can. Pictures include those of Black and White men and women, and words include those with positive and negative connotations. Completion of the online measure was estimated to take 5 minutes. The study collected the amount of time in milliseconds that it took participants to pair stimuli and analyzed it to quantify implicit bias using the central tendency measures algorithm outlined in Greenwald et al. (2003).

Statistical Analyses

A mixed-design analysis of variance (ANOVA) was constructed to measure the main effects of labeled race and speaker race on the outcome of CAPE-V severity ratings. This addressed whether labeled race had a statistically significant impact on dysphonia ratings regardless of true speaker race. Listener and speaker were entered as random variables. The interaction between labeled race and speaker race was also added to the model. This addressed whether labeled race had a statistically significant impact on dysphonia ratings only for speakers of one underlying race. The effect size for each significant ANOVA factor was calculated as squared partial curvilinear correlations (η_p^2). Effect sizes of ~0.01 were classified as small, ~0.09 as medium, and >0.25 as large (Witte & Witte, 2009). Completion of all statistical analyses was done in *Minitab* (version 19.2020.1; Minitab LLC, State College, PA), with the criterion for statistical significance set at α=0.05.

Results

Descriptive statistics

Thirty listeners completed 40 ratings per study session (not including repeated trials for reliability), resulting in 80 ratings per listener and 2400 ratings total. As shown in Table 1, these ratings were evenly distributed among four conditions: Black speakers labeled Black, Black speakers labeled White, White speakers labeled White, and White speakers labeled Black.

Condition	N	Mean Rating	SD	Range of Ratings
Black speakers				
Labeled Black	600	21.39	19.35	0 - 78
Labeled White	600	24.48	20.20	0 - 92
White speakers				
Labeled White	600	16.81	16.63	0 - 84
Labeled Black	600	14.74	15.21	0 - 75

Table 1. CAPE-V rating descriptive information across race label conditions.

Mixed-design ANOVA

We hypothesized a main effect of race labeling, regardless of underlying speaker race. Specifically, we hypothesized that speakers would be rated as less severely dysphonic by novice raters when their race was labeled as Black. The mixed-design ANOVA revealed that there was no significant effect of speaker race ($F(1, 2367) = 3.78$, $p = .06$), such that Black speaker ratings ($M = 22.93$, $SD_{pool} = 19.78$) were not significantly different from White speaker ratings ($M = 15.78$, $SD_{pool} = 15.94$) across

conditions. There were large and significant effects of rater ($F(29, 2367) = 53.73$, $p < .001$, $\eta_p^2 = .40$) and speaker ($F(38, 2367) = 68.02$, $p < .001$, $\eta_p^2 = .53$) as expected.

As hypothesized, there was a small but significant main effect of race labeling, ($F(1, 2367) = 33.41$, $p < .001$, $\eta_p^2 = .01$), such that labeling a speaker as Black resulted in less severe dysphonia ratings (Figure 2). Black speakers were rated as significantly less severe when labeled Black ($M = 24.48$, $SD = 20.2$) compared to when labeled White ($M = 21.39$, $SD = 19.35$) and White speakers were rated as significantly less severe when labeled Black ($M = 14.74$, $SD = 15.21$) compared to when labeled White ($M = 16.81$, $SD = 16.63$).

There was no significant interaction between labeled race and speaker race ($F(1, 2367) = 1.32$, $p = .25$), meaning the main effect of race labeling was true regardless of underlying speaker race. These findings support the hypothesis that race labeling would influence CAPE-V severity ratings regardless of true speaker race, as this result was true for both White and Black speakers.

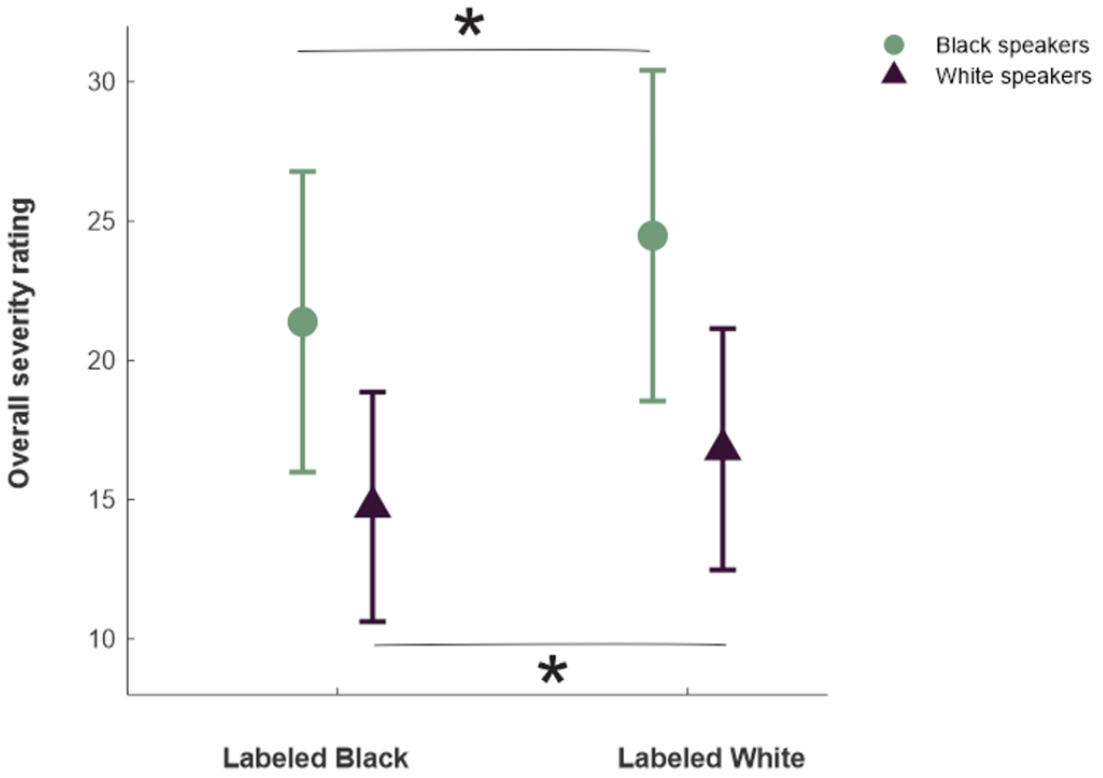

Figure 2. Mean CAPE-V severity ratings of Black and White speakers across labeling conditions. Error bars are 95% confidence intervals.

Pearson correlation

The hypothesis predicting that the difference in mean ratings (rating when labeled Black - rating when labeled White) would negatively correlate with IAT scores was tested using a Pearson correlation. A negative difference indicated that speakers labeled Black were rated as significantly less dysphonic. Positive IAT scores indicated anti-Black implicit bias. Contrary to the hypothesis, these variables were not significantly correlated ($r(28) = -.05$, $p = .78$; Figure 3).

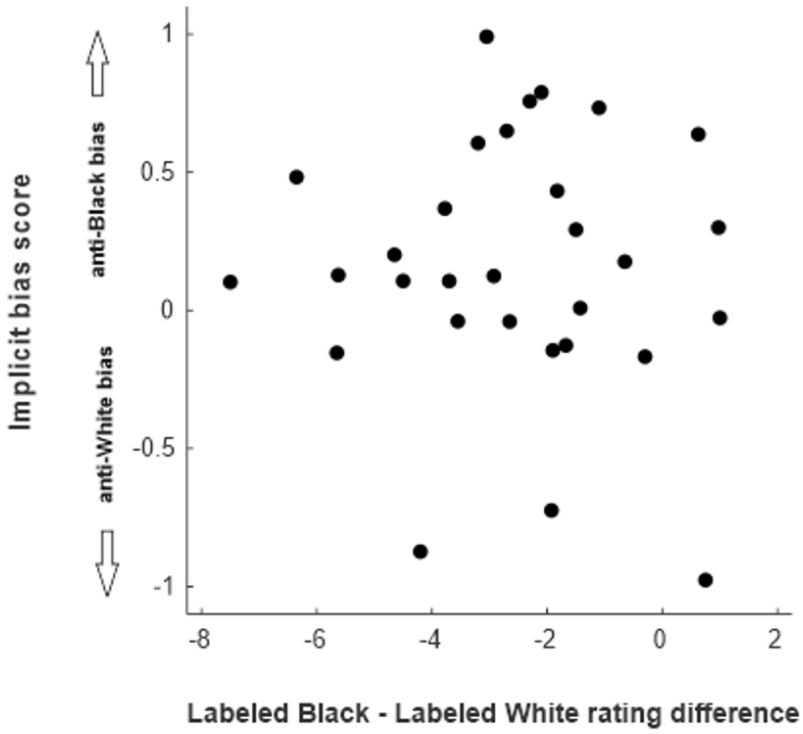

Figure 3. Listener scores on the Harvard IAT and mean differences of speaker ratings (Labeled Black - Labeled White). Negative rating differences indicate White race labeling increasing perceived severity of dysphonia.

Reliability

Interrater reliability amongst the 30 raters was calculated to be excellent at ICC = .97. Intrarater reliability was calculated to be good on the CAPE-V ratings completed twice by each rater, with the average reliability being r = .78 (SD = .15, range: .48 - .99).

Discussion

The aim of the current study was to further understand the role of implicit racial bias in auditory-perceptual evaluations of dysphonic voices. This was assessed by determining if a biasing effect existed for novice listeners in their auditory-perceptual

ratings of Black and White speakers of General American English who have voice disorders.

Impact of Race Labeling

We hypothesized that speakers would be rated as less severely dysphonic by novice listeners on the CAPE-V when their race was labeled as Black. The results from the mixed-design ANOVA support this hypothesis. Black and White speakers were both rated as significantly less severely dysphonic when their race was labeled as Black. This aligns with the minimizing bias for Black individuals presenting with various health conditions that has been shown throughout healthcare literature (Arora et al., 2018; Javitt et al., 1991; Johnson-Jennings et al., 2020; Odonkor et al., 2021; Rathore et al., 2000). This finding suggests that knowing a patient's race will impact a novice speech-language pathologist's perception of their voice disorder severity. As mentioned previously, there is no biological basis for race-based differences in voice (Sapienza, 1997; Xue & Fucci, 2000), therefore the minimizing bias for Black individuals is likely the result of implicit racial biases held by listeners. This finding provides evidence of another way that auditory-perceptual evaluations are subject to perceptual bias, in addition to previous findings supporting that medical diagnosis knowledge and accuracy of listed diagnosis can bias auditory-perceptual evaluations (Eadie et al., 2011; Sauder & Eadie, 2020).

Although this finding is significant and concerning for the care of Black patients with voice disorders, it is important to recognize the difference between statistical and clinical significance. The aforementioned results are statistically significant, but the mean difference between Black speakers labeled Black and Black speakers labeled White was

only 3.09 points on the CAPE-V, and the mean difference between White speakers labeled White and White speakers labeled Black was only 2.07 points on the CAPE-V. Severity scores for the CAPE-V range from 0-100 with anchors at 10 for mild, 35 for moderate, and 72 for severe (Kempster et al., 2009). Given this, a rating difference of 2-3 points would not likely alter the interpretation of a patient's CAPE-V severity rating, which is reassuring as it applies to the clinical care of Black patients.

Performance on IAT and CAPE-V Ratings

A negative correlation between difference in mean ratings (Labeled Black - Labeled White) and Harvard IAT scores was hypothesized and tested using a Pearson correlation. No significant relationship was found between these two variables, suggesting no trend in a listener's performance on the implicit racial bias measure compared to their performance on the CAPE-V ratings. Multiple explanations could contribute to this finding.

First, it is possible that the Harvard IAT was not a representative measure of these listeners' implicit biases. Although the Harvard IAT has been used in research for over two decades and was initially found both valid and reliable, recent studies have suggested weak predictive validity (Meissner et al., 2019) and weak construct validity (Schimmack, 2021) of the measure. Second, activation of implicit biases will differ situationally. Research shows that cognitive load (e.g., mental fatigue) can increase the likelihood of activating implicit biases that would alter task performance (Burgess, 2010; Johnson et al., 2016), and that increased subjectivity or ambiguity increases our reliance on implicit associations (Fridell, 2017; Hirsh et al., 2015). Based on this, it is possible that the IAT

had a lowered cognitive load or was less subjective for some listeners than the CAPE-V rating task, which may explain why performance on the two measures did not follow the predicted trend.

Third, the IAT is a well-known measure given its frequent use since 1998. As listeners began the task, they may have been aware that the IAT would be measuring implicit racial biases, whereas they were not aware that race labeling was being manipulated in the CAPE-V rating task. Task awareness may have influenced why implicit racial biases were activated differently on the IAT compared to on CAPE-V ratings, and this awareness may have resulted in listeners attempting to alter their performance on what was intended to be an automatic measure.

Finally, all participants of this study underwent anti-racist/anti-bias training throughout their two years in the Boston University MS-SLP program. In the program's training, students attended three lectures on intersectionality, bias, and privilege that highlighted action steps towards becoming an anti-racist/anti-bias speech-language pathologist. Students each received a workbook with written reinforcement of lecture content and additional resources for continuing education. Boston University's MS-SLP program also has a student-run antiracist focus group that has collaborated with faculty to increase discussion on ways to be anti-racist and anti-biased across course content areas (e.g., Evaluation and Diagnosis, Speech Sound Disorders, Aphasia). Questions students were tasked with answering during their training included "Have you ever caught yourself in a thought that was driven by implicit bias? What happened" and "How is my privilege showing up in the assumptions that I make about the individuals with whom I

work?" Similarly, all students were required to read a text on racial disparities in healthcare for the program's Healthcare Seminar course. This consistent emphasis on implicit bias retraining woven throughout the program may have influenced this particular group of listeners' performance in the study, as humans have the capacity to alter their implicit biases with intentional retraining efforts after gaining awareness of biases (Devine et al., 2012; FitzGerald et al., 2019; Liu et al., 2022).

Future Directions for Clinical Practice

Implicit biases permeate through all aspects of our social interactions, but they can be altered through conscious efforts made to unlearn the associations our brains have automatically built over time. An initial step in one's conscious efforts would likely be taking implicit bias tests to increase awareness of one's automatic beliefs, especially those that contradict their outward beliefs. There is a growing body of evidence suggesting that implicit bias reduction strategies are successful. Research has shown significant reductions in implicit racial bias specifically as a result of a 12-week intervention (Devine et al., 2012). Given the evidence of cognitive load (e.g., distractions and time pressures) increasing implicit bias activation, healthcare settings have been recommended to increase staffing to decrease productivity demand, decrease unnecessary information in electronic medical records, and increase use of cognitive aids such as algorithms, checklists, and clinical decision support tools (Burgess, 2010). Mindfulness has been suggested as an implicit bias retraining intervention for clinicians to increase provider awareness and control over implicit biases towards patients (Burgess et al., 2017). Other studies have suggested that intergroup exposure, or even just imagining intergroup

contact, can increase positive implicit attitudes towards that group (Turner & Crisp, 2010). When our exposure to a group is limited, we rely on stereotypes, but increased exposure to examples that counter the stereotypes we rely on may help as well. A concern many individuals have with taking the IAT is discomfort and fear of what their responses may reveal, therefore increased discussion around the inevitable nature of implicit biases may facilitate decreased stigma around this topic.

The growing body of research on implicit bias training is promising, but it is important to recognize that implicit biases are built up over one's lifetime, and increased exposure to information that confirms our biases serves to strengthen the negative associations we hold. As a result, retraining implicit biases cannot happen over a short period of time, and it does not happen without concentrated and intentional efforts. Although a devoted day of implicit bias training in an academic setting or work environment could potentially increase awareness of implicit biases, acceptance of their inevitable nature, and understanding of how they may contradict our stated beliefs, one concentrated effort will not override the automatic processes that have been reinforced over a lifetime. Further evidence is needed on the most effective methods of implicit bias retraining, but it is encouraging to know that increased awareness of implicit biases and intention to change them has been found effective.

Finally, it is recommended that voice-specialized clinicians recognize the CAPE-V as a subjective tool. Cross-checking evaluations with colleagues and doing team calibrations may prioritize patient care and reduce discrepancies between providers. Ultimately, improved patient care is facilitated through recognition that we cannot be

fully objective in evaluating patients. Implicit biases are pervasive, but taking conscious efforts to retrain our negative biases may improve clinical judgment and decision-making.

Limitations and Future Study Directions

A limitation of this study is that all listeners were novice clinicians. Beyond this, most intended to be generalist clinicians or specialists in content areas not including voice disorders upon completion of their graduate training. Given these listeners were less experienced with CAPE-V ratings and less likely to use them in future clinical settings, a future direction for this research would be conducting this study with experienced voice clinicians who are more likely to use the CAPE-V on a regular basis in qualifying the severity of voice disorders.

Another limitation is that the speakers of this study were skewed as milder in their voice disorder severity as measured by an expert clinician who conducted CAPE-V ratings. As mentioned previously, the CAPE-V severity scores range from 0-100 with anchors at 10 for mild, 35 for moderate, and 72 for severe (Kempster et al., 2009). Based on the voice-specialized clinician ratings, the average severity of all Black speakers was 20.9 out of 100 (*SD* = 12.9, range = 4.8 - 50.7), and the average severity of all White speakers was 20.0 out of 100 (*SD* = 17.4, range = 3.4 - 82.0). Although moderate and severe voice disorders were included in the sample, the majority of speakers in the study had mild voice disorders. A future direction may include replicating the study with more moderate and severe voices to see if increasing the range of dysphonia severity will influence the findings.

It is important to note the narrowness of this study's scope as it focused exclusively on auditory-perceptual voice evaluations. Although these particular findings are unlikely to be clinically significant, implicit racial bias permeates through all aspects of healthcare practice. Auditory-perceptual evaluations such as the CAPE-V are a frequently used tool in clinical voice evaluations, but they are used in combination with other measures, including (but not limited to) case history interviews, patient-reported measures, acoustic and aerodynamic data, and stroboscopic findings to comprehensively evaluate and diagnose voice disorder presence and severity. Future research is needed to address the impact of implicit racial bias on these measures given the existing evidence that implicit racial bias may cause Black patients to be viewed as less cooperative during medical visits, more susceptible to negative interactions with providers, and overall less likely to receive standard quality of medical care (Arora et al., 2018; FitzGerald & Hurst, 2017; Green et al., 2007; Maina et al., 2018; Rathore et al., 2000; Sabin & Greenwald, 2012).

Conclusion

This study found that both Black and White speakers were rated as less severe on the CAPE-V when their race was labeled as Black. This suggests a minimizing bias for Black patients with voice disorders who present to speech-language pathologists for evaluation. This finding is statistically significant, but it is likely not clinically significant. These results will contribute to the understanding of how demographic information unrelated to a voice disorder may impact a patient's visit with a clinician. Future directions include replicating this study with more experienced clinicians and with

more severe speakers. Further research is needed to determine the most effective interventions for implicit bias retraining and the additional ways that implicit racial bias impacts comprehensive voice evaluations.

BIBLIOGRAPHY

American Speech-Language-Hearing Association. (1993). *Definitions of communication disorders and variations* [Relevant paper]. https://www.asha.org/policy/RP1993-00208/

American Speech-Language-Hearing Association. (2019). High- lights and trends: Member and affiliate counts, year end 2018. https://www.asha.org/siteassets/surveys/2010-2019-member- and-affiliate-profiles.pdf

American Speech-Language-Hearing Association. (2020). Response to racism [Position statement]. http://www.asha.org/policy

Arora, S., Stouffer, G. A., Kucharska-Newton, A., Vaduganathan, M., Qamar, A., Matsushita, K., Kolte, D., Reynolds, H. R., Bangalore, S., Rosamond, W. D., Bhatt, D. L., & Caughey, M. C. (2018). Fifteen-Year Trends in Management and Outcomes of Non–ST-Segment–Elevation Myocardial Infarction Among Black and White Patients: The ARIC Community Surveillance Study, 2000–2014. *Journal of the American Heart Association*, *7*(19), e010203. https://doi.org/10.1161/JAHA.118.010203

Bainbridge, K. E., Roy, N., Losonczy, K. G., Hoffman, H. J., & Cohen, S. M. (2017). Voice disorders and associated risk markers among young adults in the United States. *Laryngoscope*, *127*(9), 2093–2099. https://doi.org/10.1002/lary.26465

Bhattacharyya, N. (2014). The prevalence of voice problems among adults in the United States. *Laryngoscope*, *124*(10), 2359–2362. https://doi.org/10.1002/lary.24740

Bureau of Professional Licensing. (2021). *FAQs for Implicit Bias Training*. Licensing and Regulatory Affairs (LARA). https://www.michigan.gov/-/media/Project/Websites/lara/bpl/Shared-Files/Implicit-Bias-Training-FAQs.pdf?rev=3bc940b9f585408db605070baf61c83e

Burgess, D. J. (2010). Are providers more likely to contribute to healthcare disparities under high levels of cognitive load? How features of the healthcare setting may lead to biases in medical decision making. *Medical Decision Making*, *30*(2), 246-257. https://doi.org/10.1177/0272989x09341751

Burgess, D. J., Beach, M. C., & Saha, S. (2017). Mindfulness practice: A promising approach to reducing the effects of clinician implicit bias on patients. *Patient Education and Counseling*, *100*(2), 372-376. https://doi.org/https://doi.org/10.1016/j.pec.2016.09.005

Charity, A. H. (2008). African American English: An Overview. *Perspectives on Communication Disorders and Sciences in Culturally and Linguistically Diverse (CLD) Populations*, *15*(2), 33-42. https://doi.org/doi:10.1044/cds15.2.33

Dalton, S., & Villagran, M. (2018). Minimizing and addressing implicit bias in the workplace: Be proactive, part one. *College & Research Libraries News*, *79*, 478. https://doi.org/10.5860/crln.79.9.478

Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology*, *48*(6), 1267-1278. https://doi.org/10.1016/j.jesp.2012.06.003

Eadie, T., Sroka, A., Wright, D. R., & Merati, A. (2011). Does knowledge of medical diagnosis bias auditory-perceptual judgments of dysphonia? *Journal of Voice*, *25*(4), 420-429. https://doi.org/10.1016/j.jvoice.2009.12.009

Easton, C., & Verdon, S. (2021). The Influence of Linguistic Bias Upon Speech-Language Pathologists’ Attitudes Toward Clinical Scenarios Involving Nonstandard Dialects of English. *American Journal of Speech-Language Pathology*, *30*(5), 1973-1989. https://doi.org/doi:10.1044/2021_AJSLP-20-00382

Ellis, C., Jacobs, M., & Kendall, D. (2021). The Impact of Racism, Power, Privilege, and Positionality on Communication Sciences and Disorders Research: Time to Reconceptualize and Seek a Pathway to Equity. *American Journal of Speech-Language Pathology*, *30*(5), 2032-2039. https://doi.org/doi:10.1044/2021_AJSLP-20-00346

Evans, K. E., Munson, B., & Edwards, J. (2018). Does Speaker Race Affect the Assessment of Children's Speech Accuracy? A Comparison of Speech-Language Pathologists and Clinically Untrained Listeners. *Language, Speech, and Hearing Services in Schools*, *49*(4), 906-921. https://doi.org/10.1044/2018_lshss-17-0120

Fairbanks, G. (1960). Voice and articulation drillbook, 2nd edn. New York: Harper & Row. pp124-139.

Farrugia-Bernard, A. M. (2018). Speech-language pathologists as determiners of the human right to diversity in communication for school children in the US. *International Journal of Speech-Language Pathology*, *20*(1), 170-173. https://doi.org/10.1080/17549507.2018.1406002

Ferrand, Carole T. Voice Disorders: Scope of Theory and Practice. Boston: Pearson, 2012. Print.

FitzGerald, C., & Hurst, S. (2017). Implicit bias in healthcare professionals: a systematic review. *BMC Medical Ethics*, *18*(1), 19. https://doi.org/10.1186/s12910-017-0179-8

FitzGerald, C., Martin, A., Berner, D., & Hurst, S. (2019). Interventions designed to reduce implicit prejudices and implicit stereotypes in real world contexts: a systematic review. *BMC Psychology*, *7*(1), 29. https://doi.org/10.1186/s40359-019-0299-7

Fridell, L. A. (2017). The Science of Implicit Bias and Implications for Policing. In L. A. Fridell (Ed.), *Producing Bias-Free Policing: A Science-Based Approach* (pp. 7-30). Springer International Publishing. https://doi.org/10.1007/978-3-319-33175-1_2

Green, A. R., Carney, D. R., Pallin, D. J., Ngo, L. H., Raymond, K. L., Iezzoni, L. I., & Banaji, M. R. (2007). Implicit bias among physicians and its prediction of thrombolysis decisions for black and white patients. *Journal of General Internal Medicine*, *22*(9), 1231-1238. https://doi.org/10.1007/s11606-007-0258-5

Greenwald, T., Banaji, M., & Nosek, B. (1998). Race Implicit Association Test. *Project Implicit*. https://implicit.harvard.edu/implicit/

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*(2), 197-216. https://doi.org/10.1037/0022-3514.85.2.197

Helou, L. B., Solomon, N. P., Henry, L. R., Coppit, G. L., Howard, R. S., & Stojadinovic, A. (2010). The Role of Listener Experience on Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) Ratings of Postthyroidectomy Voice. *American Journal of Speech-Language Pathology*, *19*(3), 248-258. https://doi.org/doi:10.1044/1058-0360(2010/09-0012)

Hirsh, A. T., Hollingshead, N. A., Ashburn-Nardo, L., & Kroenke, K. (2015). The Interaction of Patient Race, Provider Bias, and Clinical Ambiguity on Pain Management Decisions. *The Journal of Pain*, *16*(6), 558-568. https://doi.org/https://doi.org/10.1016/j.jpain.2015.03.003

Issues in Ethics: Cultural and Linguistic Competence. (2017). American Speech-Language-Hearing Association; American Speech-Language-Hearing Association. Retrieved March 20, 2022, from https://www.asha.org/practice/ethics/cultural-and-linguistic-competence/

Javitt, J. C., McBean, A. M., Nicholson, G. A., Babish, J. D., Warren, J. L., & Krakauer, H. (1991). Undertreatment of Glaucoma among Black Americans. *New England*

Journal of Medicine, *325*(20), 1418-1422. https://doi.org/10.1056/nejm199111143252005

Johnson, T. J., Hickey, R. W., Switzer, G. E., Miller, E., Winger, D. G., Nguyen, M., Saladino, R. A., & Hausmann, L. R. (2016). The Impact of Cognitive Stressors in the Emergency Department on Physician Implicit Racial Bias. *Academic Emergency Medicine*, *23*(3), 297-305. https://doi.org/10.1111/acem.12901

Johnson-Jennings, M., Duran, B., Hakes, J., Paffrath, A., & Little, M. M. (2020). The influence of undertreated chronic pain in a national survey: Prescription medication misuse among American indians, Asian Pacific Islanders, Blacks, Hispanics and whites. *SSM - Population Health*, *11*, 100563. https://doi.org/https://doi.org/10.1016/j.ssmph.2020.100563

Kempster, G. B., Gerratt, B. R., Verdolini Abbott, K., Barkmeier-Kraemer, J., & Hillman, R. E. (2009). Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. *American Journal of Speech-Language Pathology*, *18*(2), 124-132. https://doi.org/10.1044/1058-0360(2008/08-0017)

Kester, E. S. (2014). *Difference or Disorder? Understanding Speech and Language Patterns in Culturally and Linguistically Diverse Students* (First Edition). Bilinguistics, Inc.

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal Chiropractic Medicine*, *15*(2), 155-163. https://doi.org/10.1016/j.jcm.2016.02.012

Liu, F. F., Coifman, J., McRee, E., Stone, J., Law, A., Gaias, L., Reyes, R., Lai, C. K., Blair, I. V., Yu, C.-l., Cook, H., & Lyon, A. R. (2022). A Brief Online Implicit Bias Intervention for School Mental Health Clinicians. *International Journal of Environmental Research and Public Health*, *19*(2), 679. https://www.mdpi.com/1660-4601/19/2/679

Mahendra, N., & Spicer, J. (2014). Access to Speech-Language Pathology Services for African-American Clients with Aphasia: A Qualitative Study. *Perspectives on Communication Disorders and Sciences in Culturally and Linguistically Diverse (CLD) Populations*, *21*(2), 53-62. https://doi.org/10.1044/cds21.2.53

Maina, I. W., Belton, T. D., Ginzberg, S., Singh, A., & Johnson, T. J. (2018). A decade of studying implicit racial/ethnic bias in healthcare providers using the implicit association test. *Social Science & Medicine*, *199*, 219-229. https://doi.org/10.1016/j.socscimed.2017.05.009

Meissner, F., Grigutsch, L. A., Koranyi, N., Müller, F., & Rothermund, K. (2019). Predicting Behavior With Implicit Measures: Disillusioning Findings, Reasonable Explanations, and Sophisticated Solutions [Review]. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02483

Milne, A. E., Bianco, R., Poole, K. C., Zhao, S., Oxenham, A. J., Billig, A. J., & Chait, M. (2021). An online headphone screening test based on dichotic pitch. *Behavior Research Methods*, *53*(4), 1551-1562. https://doi.org/10.3758/s13428-020-01514-0

Morgan, P. L., Farkas, G., Hillemeier, M. M., Li, H., Pun, W. H., & Cook, M. (2017). Cross-Cohort Evidence of Disparities in Service Receipt for Speech or Language Impairments. *Exceptional Children*, *84*(1), 27-41. https://doi.org/10.1177/0014402917718341

Mulchan, S. S., Wakefield, E. O., & Santos, M. (2021). What COVID-19 Teaches Us About Implicit Bias in Pediatric Health Care. *Journal of Pediatric Psychology*, *46*(2), 138-143. https://doi.org/10.1093/jpepsy/jsaa131

Odonkor, C. A., Esparza, R., Flores, L. E., Verduzco-Gutierrez, M., Escalon, M. X., Solinsky, R., & Silver, J. K. (2021). Disparities in Health Care for Black Patients in Physical Medicine and Rehabilitation in the United States: A Narrative Review. *PM & R*, *13*(2), 180-203. https://doi.org/10.1002/pmrj.12509

Rathore, S. S., Lenert, L. A., Weinfurt, K. P., Tinoco, A., Taleghani, C. K., Harless, W., & Schulman, K. A. (2000). The effects of patient sex and race on medical students's ratings of quality of life. *The American Journal of Medicine*, *108*(7), 561-566. https://doi.org/10.1016/S0002-9343(00)00352-1

Rubin, D. L. (1992). Nonlanguage Factors Affecting Undergraduates' Judgments of Nonnative English-Speaking Teaching Assistants. *Research in Higher Education*, *33*(4), 511-531. http://www.jstor.org/stable/40196047

Sabin, J. A., & Greenwald, A. G. (2012). The influence of implicit bias on treatment recommendations for 4 common pediatric conditions: pain, urinary tract infection, attention deficit hyperactivity disorder, and asthma. *American Journal of Public Health*, *102*(5), 988-995. https://doi.org/10.2105/ajph.2011.300621

Sapienza, C. M. (1997). Aerodynamic and acoustic characteristics of the adult African American voice. *Journal of Voice*, *11 4*, 410-416.

Sauder, C., & Eadie, T. (2020). Does the Accuracy of Medical Diagnoses Affect Novice Listeners' Auditory-Perceptual Judgments of Dysphonia Severity? *Journal of Voice*, *34*(2), 197-207. https://doi.org/10.1016/j.jvoice.2018.08.001

Schimmack, U. (2021). Invalid Claims About the Validity of Implicit Association Tests by Prisoners of the Implicit Social-Cognition Paradigm. *Perspectives on Psychological Sciences*, *16*(2), 435-442. https://doi.org/10.1177/1745691621991860

Smedley, B. D., Stith, A. Y., & Nelson, A. R. (2003). *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*. The National Academies Press.

Turner, R. N., & Crisp, R. J. (2010). Imagining intergroup contact reduces implicit prejudice [https://doi.org/10.1348/014466609X419901]. *British Journal of Social Psychology*, *49*(1), 129-142. https://doi.org/https://doi.org/10.1348/014466609X419901

U.S. Census Bureau QuickFacts: United States. (2021). Retrieved March 31, 2022, from https://www.census.gov/quickfacts/US

Witte, R. S., & Witte, J. S. (2009). Statistics (9th ed.). Wiley.

Xue, S. A., & Fucci, D. (2000). Effects of Race and Sex on Acoustic Features of Voice Analysis. *Perceptual and Motor Skills*, *91*(3), 951-958. https://doi.org/10.2466/pms.2000.91.3.951

Zraick, R. I., Kempster, G. B., Connor, N. P., Thibeault, S., Klaben, B. K., Bursac, Z., Thrush, C. R., & Glaze, L. E. (2011). Establishing validity of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). *American Journal of Speech- Language Pathology, 20*(1), 14-22. https://doi.org/10.1044/1058-0360(2010/09-0105)

CURRICULUM VITAE