ASHA American Speech-Language-Hearing Association

AJSLP

Research Article

Are Auditory-Perceptual Evaluations of Dysphonia by Experienced Voice Clinicians Affected by Knowledge of Speaker Race?

Daniel P. Buckley,[a,b] Kimberly L. Dahl,[a] Rachel L. Norotsky,[a] Cara Sauder,[c] Lauren F. Tracy,[b] and Cara E. Stepp[a,d]

[a] Department of Speech, Language, and Hearing Sciences, Boston University, MA [b] Department of Otolaryngology—Head and Neck Surgery, Boston University Chobanian and Avedisian School of Medicine, MA [c] Department of Speech and Hearing Sciences, University of Washington, Seattle [d] Department of Biomedical Engineering, Boston University, MA

ARTICLE INFO

Article History:
Received August 20, 2024
Revision received November 12, 2024
Accepted January 30, 2025

Editor-in-Chief: Rita R. Patel
Editor: Robin A. Samlan

https://doi.org/10.1044/2025_AJSLP-24-00355

ABSTRACT

Purpose: The purpose of this study was to investigate implicit racial bias in auditory-perceptual evaluations of dysphonic voices completed by experienced voice-focused speech-language pathologists (SLPs).
Method: Thirty SLPs specializing in voice disorders listened to audio files of 20 Black speakers and 20 White speakers of General American English with voice disorders. The SLPs rated the overall severity of dysphonia (OS) of each voice heard using a 100-unit visual analog scale and then completed the Harvard Implicit Association Test (IAT) to estimate their implicit racial bias.
Results: There were no significant main effects of speaker race or labeled race on OS ratings; however, there was a small but significant interaction effect between them: Race labeling resulted in a minimizing effect for Black speakers, but not White speakers. No significant relationship was found between Harvard IAT scores and differences in OS ratings by race-labeling condition.
Conclusions: These findings suggest that experienced, voice-focused SLPs demonstrated a minimizing bias in their auditory-perceptual ratings of dysphonia of Black speakers. This bias is small and may not be clinically significant but, in some cases, could contribute to worse clinical care of Black people with voice disorders.

The auditory-perceptual assessment of voice is a core component of voice evaluations and therapeutic intervention. Those diagnosed with dysphonia present with atypical or aberrant auditory-perceptual features in their voice quality, pitch, or resonance (Stachler et al., 2018). A multidisciplinary team of otolaryngologists and speech-language pathologists (SLPs) help to characterize this dysphonia while considering what voice quality might be expected for an individual's age, sex, gender, cultural background, or culture (Boone et al., 2014). Considering the subjective nature of auditory-perceptual evaluations and impressions of voice quality, it is unsurprising that it has been demonstrated that certain features beyond these components may influence an individual's perception of voice quality. Eadie et al. (2011) found that a bias, or the systematic deviation from a normative response (Kahneman et al., 1982), was present when listeners were presented with a speaker's diagnosis: both experienced and inexperienced individuals rated voices as more severely dysphonic. Sauder and Eadie (2020) found a similar overpathologizing bias: When novice raters were presented with inaccurate diagnoses, they produced higher ratings of roughness and breathiness. However, when the listeners were told that these individuals did not have a voice disorder, they demonstrated a minimizing bias by producing reduced ratings of overall severity of dysphonia (OS) for individuals with voice

Correspondence to Daniel P. Buckley: buckley@bu.edu. ***Disclosure:*** *Cara E. Stepp has received consulting fees from Altec, Inc./Delsys, Inc., companies focused on developing and commercializing technologies related to human movement. Stepp's interests were reviewed and are managed by Boston University in accordance with their conflict of interest policies. The other authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

disorders. These findings suggest that bias may impact both novice and experienced raters' auditory-perceptual ratings of voice quality.

Bias in health care, particularly racial bias, is an increasingly prominent topic of recent discussion. Racial bias can be described as the subconscious negative attitudes against a specific group of people based on their racial group (FitzGerald & Hurst, 2017). Implicit bias specifically relates to subconscious feelings, attitudes, or prejudice, whereas explicit racial bias references conscious beliefs or opinions held against a particular racial group or identity (A. Nelson, 2002). A systematic review examining general implicit bias found significant levels of implicit bias in health care professionals, with a negative relationship between the quality of care and the degree of implicit bias (FitzGerald & Hurst, 2017). Likewise, another systematic review demonstrated that increased implicit bias led to worse interactions between providers and their patients (Maina et al., 2018). Black patients are 2.54 times more likely to have at least one negative descriptor (e.g., noncompliant, drug-seeking) in their medical records than White patients, even after adjusting for sociodemographic and negative health characteristics (Sun et al., 2022). This is consistent with the finding that physicians view Black patients more negatively than White patients on parameters, such as adherence to medical advice and intelligence, and spend less time with them during their clinical encounters (Van Ryn & Burke, 2000). Even when physicians deny any explicit racial bias, implicit measures demonstrate a strong preference for White individuals, with implicit stereotypes that Black patients are less cooperative during medical visits (Green et al., 2007). Thus, the presence of implicit racial bias alone appears to directly impact delivered health care. For instance, Black patients are less likely to receive commonly prescribed medications and treatment protocols in the hospital setting after a heart attack (Arora et al., 2018). Similarly, it was demonstrated that medical students were less likely to diagnose Black women with angina when they presented with identical set of symptoms to those of White women (Centola et al., 2021; Rathore et al., 2000). This underestimation of the severity of conditions and underutilization of interventions by health care providers demonstrate a minimizing bias. Similar minimizing biases have been demonstrated in the literature, resulting in reduced treatment and quality of provided care for Black individuals or people of color (Centola et al., 2021; Javitt et al., 1991; Johnson-Jennings et al., 2020; Odonkor et al., 2021).

Despite the well-demonstrated impact of racial bias in health care at large, investigations of racial biases related to speech-language pathology services are more limited. In general, it has been suggested that the disproportional representation of Black students receiving special education services is multifactorial, with reduced cultural competence and responsiveness in education being suggested as a contributing factor (Othman, 2018). Papandrea et al. (2023) observed that recommendations from SLPs, psychologists, and school social workers demonstrated racial bias when assigning eligibility for special education services, with Black students less likely to be recommended for services than White children of a similar socioeconomic status. Beyond early education and specialty services, when applying to speech-language pathology graduate programs, students from underrepresented backgrounds reported needing to apply to more graduate programs than their peers (Lugo et al., 2023). These data demonstrate heightened barriers for Black individuals, from early education environments through applying for graduate training to be SLPs themselves.

Minimizing biases have been directly observed to impact the quality of speech-language pathology care. Black individuals are less likely to receive services for early intervention (Morgan et al., 2012) and specific language impairments (Morgan et al., 2017). Furthermore, Morgan et al. (2016) found that Black children were half as likely to receive SLP services compared with White children, with other studies demonstrating similar reduced access to services for speech and language disorders (Black et al., 2015; Davidson et al., 2022). Black preschoolers received significantly less alternative augmentative communication services compared to their White peers (Pope et al., 2022). In another study, when shown Black children's faces and accompanying speech sounds, both trained SLPs and untrained listeners rated productions as "more accurate" than when the stimuli were presented with accompanying White faces (Evans et al., 2018), demonstrating a minimizing bias through attributing more errors as typical for Black children. In adults, Black patients were recommended to less frequent sessions for aphasia than were White patients despite large language impairments, with 80% reporting at least one negative experience in treatment with an SLP (Mahendra & Spicer, 2014). Given the underreporting of race in SLP clinical research (Ellis, 2009; Morton & Sandage, 2022), the extent and impact of racial bias in speech-language pathology is still relatively uncharacterized. It seems likely that reduced access to care and services as well as minimizing biases within the care itself might impact clinical care in areas beyond what has already been observed in the field of speech-language pathology, such as in voice disorders.

Mayne and Namazi (2023) highlighted that social determinants of health, or personal environments that impact life functioning, quality of life (National Academies of Sciences, Engineering, and Medicine, 2019), health outcomes, and health risks, contribute to toxic stress processes that ultimately may contribute to patterns in functional voice disorders, with a disproportionate impact on more vulnerable groups such as people of color and other minority groups. Bergmark and Sedaghat (2017) highlighted the limited research into social determinants of health for various racial groups and the need to address health inequities for minority patients in the realm of otolaryngology care. As evidence of this notion, Morton and Sandage (2022) found in a meta-analysis that, of 46 voice clinical trials, only two reported race and ethnicity, and one of those underrepresented Black participants. It has been shown, however, that Black people are less likely to report voice problems than White patients (Hur et al., 2018). Furthermore, individuals who have public insurance, are part of a minoritized group, or are low-income earners are more likely to postpone care due to variables such as lack of transportation (Hur et al., 2018). Certain geographic areas have reduced access to medical SLP care in general, with these same areas representing a large proportion of low-income earners and minoritized individuals (Morton et al., 2022). However, once voice therapy is initiated, some research has found that race is not a factor in dropout rates (Hapner et al., 2009) or attendance (Misono et al., 2017).

Recently, the first prospective work examining racial bias within the realm of voice disorders was completed. Norotsky et al. (2023) investigated whether auditory-perceptual impressions of voice were impacted by race information presented to the listener. Consistent with the current consensus that race is a social rather than a biological construct (Duello et al., 2021; Smedley & Smedley, 2005), no specific acoustic differences have been demonstrated between sustained vowels produced by Black and White speakers when controlling for dialect, age, sex, height, and weight (Sapienza, 1997; Xue & Fucci, 2000). Thus, differences in auditory-perceptual ratings of White and Black speakers, when controlling for dialect, should not demonstrate specific differences. To investigate the impact of implicit racial bias, Norotsky et al. recruited graduate students studying speech-language pathology to make auditory-perceptual impressions of the OS of voices using a 100-mm Visual Analog Scale (VAS) as presented on the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V; Kempster et al., 2009). Listeners were presented with voice samples of White and Black speakers with voice disorders under two conditions: with accurate labels and with inaccurate labels of the speaker's race. Implicit bias was measured using the Harvard Implicit Association Test (IAT; Greenwald et al., n.d.). Listeners demonstrated a minimizing bias for both White and Black speakers when they were labeled as Black, reflected by reduced ratings of OS. However, despite this minimizing bias, there was no correlation between a speaker's implicit bias (measured through IAT scores) and differences in OS as a function of the race label.

It is not known if these findings would stand true in experienced raters with voice-centric training and more clinical experience and training in voice disorders. Referencing available literature, there are conflicting data regarding if experience equates to more reliability in auditory-perceptual ratings of voice (Damrose et al., 2004; Eadie & Kapsner-Smith, 2010; Eadie et al., 2007, 2010; Helou et al., 2010; Laczi et al., 2005). Considering that practicing clinicians are those delivering the majority of voice care to individuals with voice disorders, further investigation into the impact of implicit bias on auditory-perceptual evaluations of voice in experienced raters is necessary to better understand and address inequities within voice evaluations for Black patients with voice disorders.

The purpose of this study was to examine if a minimizing bias exists within the auditory-perceptual evaluations of SLPs when rating Black speakers with voice disorders. The investigation was separated into two aims: first, to investigate the impact of race labeling on auditory-perceptual ratings of OS in experienced, voice-focused SLPs; and second, to examine if differences between OS ratings associate with the strength of the SLPs' implicit racial biases. These research questions were assessed by replicating the methodology by Norotsky et al. (2023) but within a sample of voice-focused SLPs. In this study, the listeners were asked to rate voice samples and were provided with both accurate and inaccurate labels of race in a sample of age- and sex-matched speakers with voice disorders matched for OS. Implicit bias was estimated using the IAT. Our first hypothesis was that speakers labeled as Black would be rated less severely dysphonic than when they were labeled as White (i.e., a minimizing bias), considering the similar minimizing bias observed across health care and as demonstrated in Norotsky et al. Since Norotsky et al. did not detect a correlation between minimizing bias and IAT scores in novice raters, we anticipated a similar result but tested the alternative hypothesis that differences in OS ratings as a function of race label would be associated with the strength of the implicit bias (i.e., those with higher implicit bias scores would indicate a greater minimizing bias on their OS ratings).

Method

Speakers

Speakers were chosen from an existing research database at the Stepp Lab for Sensorimotor Rehabilitation Engineering at Boston University. These voice recordings comprised the same set that was used in previous research by Norotsky et al. (2023). This sample was created to include Black and White speakers who spoke General

American English (GAE). Those speaking African American English (AAE) were not included in this sample in order to eliminate potential cues as to the speaker's race. To construct this list of speakers, two speech-language pathology student research assistants were provided with a list of features of AAE that are outlined in detail in Charity (2008). This included specific lexical, phonological, syntactic, prosodic, and pragmatic features. Using these parameters, the two raters independently classified the speech of 65 speakers as GAE or AAE. Agreement between raters was achieved for 85% of the samples. Fifteen percent of samples without agreement were confirmed by a third rater, a licensed speech-language pathologist, whose classifications were used to resolve the disagreements. This process yielded two groups of speakers of GAE: one group with Black speakers and one with White speakers. All speakers in this set were diagnosed with voice disorders (see Table 1), were age- and sex-matched, and were matched for OS via CAPE-V (Kempster et al., 2009) as rated by a blinded SLP specializing in voice disorders. The rationale for this process (i.e., using one SLP's ratings to match speakers) was to create a set of speakers for whom gross differences were minimized so that interpretation of perceived differences would be related to the effect of race labeling. Furthermore, the study design (see Experimental Conditions section) allowed for individual speakers to act as their own controls, reducing importance of cross-group similarities. The sample consisted of 20 Black speakers (17 females, three males, *M* = 40.1 years, *SD* = 16.7 years, range: 18–67 years) and 20 White speakers (17 females, three males, *M* = 39.6 years, *SD* = 16.3 years, range: 19–66 years). Gender information was not collected for all participants. The average OS of White speakers was 17.6 (*SD* = 10.0, range: 3.4–33.3), and the average of Black speakers was 20.9 (*SD* = 12.6, range: 4.9–50.7).

Voice Recordings

Speakers were recorded producing the sustained vowels /ɑ/ and /i/, and the second and third sentences of 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). All audio files were recorded either in a quiet clinic room at Boston Medical Center or in a sound-treated booth at Boston University using a headset microphone (Shure Model SM35XLR, Shure Model WH20XLR, or Sennheiser Model PC131) and sampling at 44.1 kHz and 16-bit resolution. Microphones were located 7 cm from the mouth's corner at a 45° angle (Patel et al., 2018). The vowels and portion of The Rainbow Passage were combined into one speech sample per speaker and normalized for peak intensity using a custom MATLAB script (Version 9.10.0.1613233 [R2021a]; The MathWorks, Inc.).

Table 1. Laryngeal diagnoses of Black and White speakers.

Diagnosis	*n*
Vocal hyperfunction	
Phonotraumatic	12
Nonphonotraumatic	24
Vocal-fold scar	1
Paradoxical vocal-fold movement	1
Reinke's edema	1
Vocal-fold atrophy	1

Note. Total *N* = 40.

Listeners

Thirty SLPs (24 cisgender females, four cisgender males, and two nonbinary individuals, *M* = 35.2 years, *SD* = 8.4, range: 26–58; all White) participated as listeners. Listeners were all clinicians with at least 1 year of clinical practice in the area of voice disorders (i.e., 50% or more of their caseload consisting of patients with voice disorders), with an average of 5.9 years of clinical experience (*SD* = 5.3, range: 1–25) and an average of 5.1 years of clinical practice with a voice caseload 50% or greater (*SD* = 4.5 years, range: 1–19). The listeners were all speakers of English who reported typical hearing, except one listener who reported tinnitus and mild high-frequency hearing loss. Listeners were financially compensated for their time. They were told they would be completing CAPE-V ratings but were, otherwise, blinded to the purpose of the study. Investigators obtained informed consent from all listeners prior to study initiation in compliance with the Boston University Institutional Review Board (#2625).

Experimental Conditions

Listeners were randomly assigned to one of two conditions. In the first condition, all samples were first presented with the speaker's race accurately labeled, followed by an optional 10-min break, and then all samples were presented again but with inaccurate race labeling. The second condition was the inverse: Samples were first presented with inaccurate race labels, followed by the optional 10-min break, and then all samples were presented with accurate race labeling. The order of accurate–inaccurate sessions was counterbalanced. All 96 samples were rated (40 voices with 20% of audio files repeated per both conditions for intrarater reliability) in one session, which took on an average of 49.5 min (range: 38.7–76.0).

Listening Procedure

The listening procedure was completed remotely for all participants. All participants connected with a study author or research assistant on a teleconferencing platform and were instructed to use headphones throughout the study. The researcher remained present on the call, with their video and audio functions off, to be available for any technological errors. They first guided the participants to Gorilla Experiment Builder (gorilla.sc) where the experiment was hosted. All participants were instructed to adjust their system volume to a comfortable level during a sustained 4-s clip of white noise. After a comfortable volume was set, listeners completed a headphone screening, which consisted of a Huggin's pitch task in which they identified a tone present in one of three white noise segments. Successful completion confirmed that the participants were using headphones for the listening procedure (Milne et al., 2021). Some participants' headphones did not initially pass this task, and they were instructed to attempt again with other headphones. All participants ultimately passed this check before proceeding.

After passing the headphone screening, the listeners began rating their first set of voice recordings. Participants first rated 48 voices (20 unique audio files of Black speakers, 20 unique audio files of White speakers, with eight of the total 40 audio files repeated for intrarater reliability) presented in a random order. The listeners were presented with a custom graphic user interface (GUI; see Figure 1) that presented the speaker's age, sex, and their labeled race for each rating. The GUI also contained a play button and a next button for moving to subsequent samples. Listeners rated the OS of each audio sample using a VAS underneath the speaker's age, sex, and race, which ranged from 0 to 100 with an anchor placed at 10 (*mild*), 35 (*moderate*), and 72 (*severe*) in accordance with the anchor positions on the printable CAPE-V form available on the American Speech-Language-Hearing Association website (www.asha.org/form/cape-v/). Listeners were instructed to drag a bar along the VAS to the level that they felt was representative of the overall severity of the speaker's dysphonia, representing their global impression of perceptually deviant voice features. Listeners were instructed to listen to all samples at least one time but were not allowed to play a sample more than twice. After participants rated the samples in the first condition, they were given the option of taking up to a 10-min break. Following this break, the samples in the second condition were presented in a random order, with opposite race labels. Interrater reliability was calculated using an intraclass correlation coefficient (ICC [C,1]), and intrarater reliability was calculated for each listener using Pearson's correlations (r).

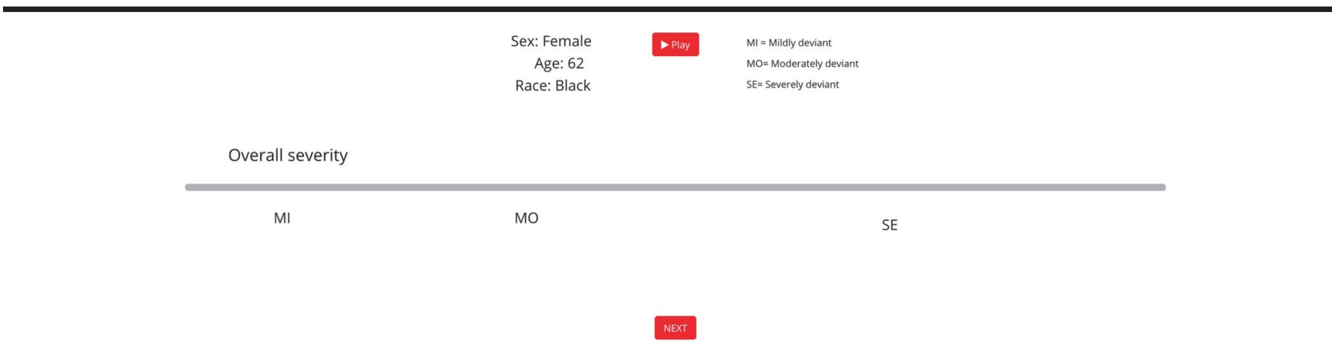
Implicit Bias Task

After completion of the CAPE-V ratings, the participants completed the Harvard IAT within the Gorilla platform to quantify their implicit racial bias. Listeners were instructed to sort presented pictures (White or Black faces) and words (good connotation or bad connotation) into left or right bins as fast as possible. The bins alternated in blocks from either collecting good connotation words with Black faces or bad connotation words with White faces, to good connotation words with White faces or bad connotation words with Black faces. The platform collected the time (in milliseconds) that it took to sort words and pictures into their respective categories to quantify implicit racial bias using the central tendency measures algorithm, which is outlined in Greenwald et al. (2003). For example, if an individual took more time to sort good connotation words into bins with Black faces than they did good connotation words with White faces, their IAT score would indicate an anti-Black bias. Positive IAT scores indicate an anti-Black bias, and negative scores indicate an anti-White bias.

Statistical Analysis

As in Norotsky et al. (2023), a mixed-design analysis of variance (ANOVA) was completed to assess the effects of labeled race, speakers' actual race, and their interaction

Figure 1. The graphic user interface presented to the listeners. The sex, age, and race are shown as an example but varied per audio sample.



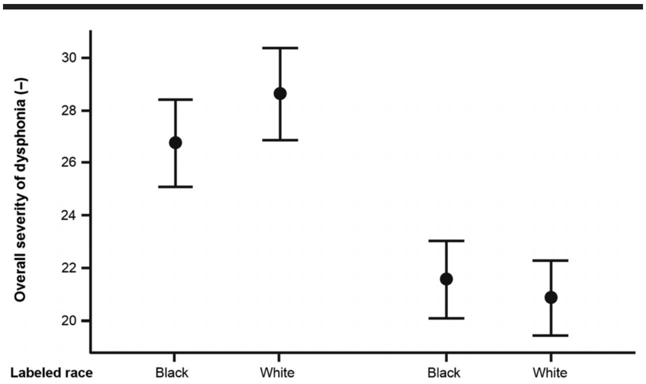
on the auditory-perceptual ratings of the OS. Speaker and listener were also included as random factors. The effect size for each statistically significant factor was calculated using the squared partial curvilinear correlation (η_p^2). Effect sizes of ~0.01 were classified as small, ~0.09 as medium, and > 0.25 as large (Witte & Witte, 2009). All statistical analyses were completed in Minitab (Version 19.2020.1; Minitab LLC), with statistical significance set a priori to $\alpha = .05$. To test if OS rating differences by race label were associated with the strength of a listener's implicit racial bias, a Pearson's correlation was calculated between IAT scores and the mean difference in ratings (OS scores when labeled Black − labeled White).

Results

Interrater and intrarater reliability were calculated across samples. Interrater reliability was interpreted according to Koo and Li (2016), and intrarater reliability was interpreted according to Cohen (1988). Interrater reliability of the listeners was ICC = .68 (95% CI [.58, .78]) or moderate reliability. The average intrarater reliability was $r = .86$ ($SD = .06$, range: .73–.97), representing a strong relationship between repeated ratings.

The ANOVA revealed no significant main effects of speaker race, $F(1, 2329) = 2.02$, $p = .163$, or labeled race, $F(1, 2329) = 1.88$, $p = .17$. However, there was a small but significant interaction effect between speaker race and labeled race, $F(1, 2329) = 9.41$, $p = .002$, $\eta_p^2 = .004$, suggesting that the race labeling affected OS ratings for Black speakers but not White speakers (see Figure 2). As expected, there were large and significant effects of listener, $F(29, 2329) = 73.92$, $p = .0$, $\eta_p^2 = .479$, and speaker, $F(38, 2329) = 119.14$, $p = .0$, $\eta_p^2 = .660$.

Figure 2. Mean ratings of overall severity of dysphonia for all speakers for both accurate and inaccurate labeling conditions. Error bars are 95% confidence intervals derived from standard deviations by speaker averaged across listeners. The *y*-axis is truncated for improved visualization, with the full analog scale ranging from 0 to 100. Higher values indicate more severe dysphonia.



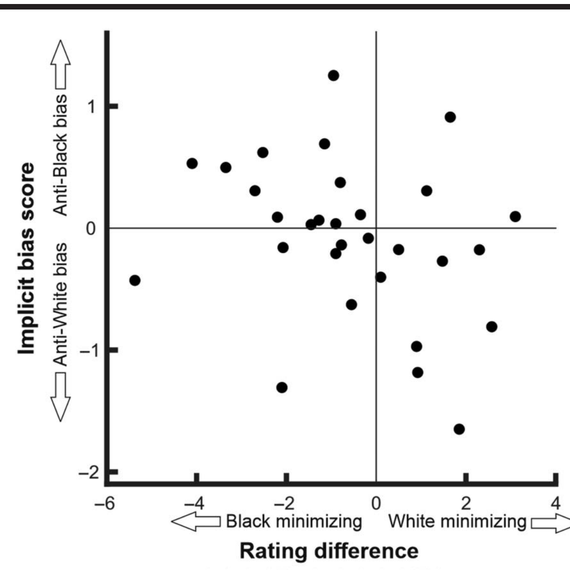
There was no significant correlation between IAT scores and individual differences in mean ratings (OS ratings when labeled BlackOS ratings when labeled White; $r(28) = -.272$, $p = .146$; see Figure 3), suggesting that the degree of anti-Black or anti-White bias did not correlate to listeners' minimizing biases of OS for Black or White speakers, which was consistent with our initial hypothesis. The potential trend in the data was opposite the alternative hypothesis tested: Those with higher anti-Black implicit bias scores, if anything, trended toward a lower degree of minimizing biases on their OS ratings.

Discussion

Impact of Race Labeling

The aim of the current work was to determine whether auditory-perceptual evaluations of dysphonia by voice-focused SLPs are affected by knowledge of speaker race. We hypothesized that speakers would be rated as less severely dysphonic when labeled as Black than when they were labeled as White, demonstrating a minimizing bias by clinicians. Although labeled race did not have a statistically significant effect on OS ratings, there was a small but significant interaction of labeled race and speaker race. Specifically, when speakers who were White were labeled as Black, there was no statistically significant

Figure 3. Listener scores on the Harvard Implicit Association Test and mean differences of speaker ratings of overall severity of dysphonia (labeled Black − labeled White). Negative rating differences indicate Black race labeling resulting in decreased perceived overall severity of dysphonia. Positive rating differences indicate Black race labeling resulting in increased overall severity of dysphonia ratings.



minimizing bias, but when Black speakers were labeled as Black, there was a minimizing bias. Black speakers labeled as Black were rated as less severely dysphonic than when they were labeled as White. This finding is in the same direction as the result in Norotsky et al. (2023); however, they found a main effect of labeled race and did not detect an interaction between labeled race and actual speaker race.

Although the statistical findings in the current study differ from Norotsky et al. (2023), these apparent differences should be interpreted with caution. The main effect of labeled race shown by the novice clinicians in Norotsky et al. was statistically significant, but with a very small effect size η_p^2 = .01. Likewise, in the current study, the interaction effect between labeled and actual speaker race was very small. Although Norotsky et al. did not detect an interaction effect, the average effect of race labeling was larger in the Black speakers than in the White speakers (3.09 mm vs. 2.07 mm).

One potential explanation for the interaction effect between labeled and actual speaker race in the current study relates to a floor effect in the OS ratings. The Black speakers on average and across race labeling conditions (i.e., the average of when labeled Black and labeled White) were judged as more severely dysphonic (*M* = 27.7, *SD* = 16.0) than the White speakers (*M* = 21.2, *SD* = 12.3). Although the mean OS was only 6.3 mm apart, this difference may have been exacerbated by the nonlinear anchor placement on the VAS. Anchors were consistent with the current printable version of the CAPE-V, with mild placed at 10 mm, moderate placed at 35 mm, and severe placed at 72 mm (see Figure 1). This nonlinear representation then may contribute to nonlinear relationships along the scale in terms of changes to auditory-perceptual impressions of dysphonia. For example, a 10-mm change from 20 to 10 mm would result in an OS marker that could be interpreted as moving from the mild–moderate to mild categories of severity. That same 10-mm change higher in the scale, such as from 45 to 35 mm, might be interpreted as less significant of a change since both remain in the moderate category. White speakers were rated as 6.3 mm less severe across race labeling conditions, with their average OS rated as 21.2 mm. In addition to being lower, more of these ratings were in the mild-to-moderate range than the ratings for the Black speakers. This might have resulted in a floor effect, particularly for the White speakers. The same speaker set used as stimuli in the current study was identical to the one used in Norotsky et al. (2023), in which the ratings by novice clinicians resulted in a similar average difference between speakers (White speakers rated as 7.0 mm less severe than Black speakers). However, the mean OS for both Black and White speakers was rated lower by those novice clinicians than by the voice-focused SLPs in the current study (22.9 and 15.8 mm for Black and White speakers, respectively), and their variability was also higher (19.8 vs. 15.9 mm for Black and White speakers, respectively). The fact that both speaker groups were more solidly in the mild-to-moderate range and the increased variability in ratings by novice clinicians may explain the lack of an interaction effect in Norotsky et al.

Although we did observe a statistically significant minimizing bias for Black speakers, the average difference in OS for Black speakers as a function of race label was only 1.85 mm. Norotsky et al. (2023) found that Black speakers labeled as Black were rated as 3.09 mm less dysphonic by novice clinicians than when labeled White, which suggests a similar minimizing bias for Black speakers but to a slightly larger degree. For reference, on the nonlinear CAPE-V, there are 25 mm between the categories mild and moderate and 37 mm between moderate and severe. Thus, the small rating differences observed in the current work as well as those found in Norotsky et al. are unlikely to have a clinically meaningful difference in most speakers with voice disorders.

Implicit Racial Bias and Ratings of Overall Severity of Dysphonia

No relationship was found between the degree of differences in OS when speakers were labeled as Black or White and the implicit racial bias scores of the listeners. Of the 30 SLP raters, 15 (50%) presented with positive IAT scores indicative of an anti-Black bias and the remaining half presented with an anti-White bias (i.e., negative scores). From the same group, 19 (63%) listeners demonstrated a minimizing bias when speakers were labeled as Black and 11 (37%) did not. Despite this range of data, no observed correlation was appreciated (see Figure 3).

The lack of relationship between implicit bias and then the degree of minimizing bias on OS ratings is consistent with two other relevant works. No correlation was detected between immigrant-specific IAT scores of school-based SLPs and their prioritization or use of best-practice evaluation procedures when assessing multilingual children (M. Nelson & Wilson, 2021). Furthermore, Norotsky et al. (2023) also did not find a correlation between implicit racial bias IAT scores of speech-language pathology students and the degree of minimizing bias on their OS ratings. It is possible that the IAT was not actually representative of the individual's actual implicit bias. Although the Harvard IAT has been used for decades as a reliable measure of implicit bias, recent studies have suggested that there might be weak construct validity (Schimmack, 2021) and weak predictive validity (Meissner et al., 2019).

One possible alternative interpretation relates to the finding that the degree of cognitive load (e.g., mental fatigue, cognitive effort) associated with a task may increase the likelihood of implicit bias impacting its performance: Higher cognitive loads may result in more observable bias (Burgess, 2010; Johnson et al., 2016). Likewise, increased ambiguity or subjectivity increases reliance on implicit associations (Fridell, 2017; Hirsh et al., 2015). Therefore, it is possible that the auditory-perceptual evaluation task might have presented low-cognitive demands for the voice-focused SLPs who rate voices on a near-daily basis. This level of heightened familiarity and low-cognitive effort in some may then explain the lack of relationship between IAT scores and minimizing bias in their ratings of voice quality. Norotsky et al. (2023) explained their absent correlation through the same notion, which increased cognitive demands activate implicit bias, although the cognitive load for the novice listeners completing the auditory-perceptual ratings was likely higher than that of the SLP raters. It may simply be the difference in cognitive load between the two tasks (i.e., the auditory-perceptual ratings vs. the IAT) that do not yield a correlation: The ability to find a relationship might be dependent upon the cognitive load of the two tasks being more closely matched. This might also explain in part why the novice clinicians in Norotsky et al. demonstrated a larger overall minimizing bias in OS ratings as compared to the experienced SLPs in the present study, due to a higher cognitive demand and reduced familiarity with the task of rating voice quality in novice clinicians.

Implications for Clinical Practice

This study detected a small minimizing bias in OS ratings of voice-focused SLPs when evaluating Black speakers. This information at face value should be a point of reflection and discussion for voice clinicians. Voice-centric clinicians may wish to identify whether this bias might exist more broadly and across other evaluation and intervention practices and to consciously reflect on this at points along their training and clinical practice. Promoting discussions around this bias may be helpful for SLPs (and speech-language pathology students) to consciously mitigate any minimizing biases for Black speakers. Incorporating direct antiracism and implicit bias training, as well as adjusting systematic structures within health care environments, may mitigate the effects of bias on health care interactions and quality of care (Cooper et al., 2022; Hagiwara et al., 2020; Marcelin et al., 2019; Sim et al., 2021). Vora et al. (2021) suggest incorporating simulation-based trainings and facilitated discussion to address racial disparities within health care to begin to dismantle frameworks upholding structural racism. It has also been demonstrated that bias in health care recommendations for Black patients can be mitigated through use of peer networks, in which a clinician engages in feedback and the exchange of clinical interpretations of a diagnosis from a large pool of other clinicians (Centola et al., 2021). Morton-Jones and Timmons Sund (2023) suggest other active measures for providing culturally responsive voice care, including self-exploration of intersectional identities, engaging with perspective taking (i.e., empathy), increasing exposure to diversity, and larger systemic shifts such as recruiting SLPs from culturally diverse backgrounds into the field and in health care settings. These active measures are likely more effective than adopting an attitude of "racial color blindness," in which a person's race is ignored or not considered: This strategy has negative impacts on speech-language pathology services and health care at large (Yu et al., 2022). Avoiding potentially uncomfortable conversations may lead to suppression of important discussions and reflection surrounding antiracist practices, which may hinder culturally responsive care. Specifically, within clinical voice training, it might be important to reemphasize that, when controlling for dialect, no differences in acoustics have been demonstrated between Black and White speakers; there should, therefore, also not be differences in a clinician's threshold for perceived dysphonia within auditory-perceptual evaluations of voice.

A final notion to consider when interpreting these findings is that this study does not account for all the possible ways in which bias can be activated. In this study, a speaker's race was conveyed on the screen through text without any visual information provided to the listener (i.e., a picture of the speaker). In a real clinical setting, it is possible that bias may be activated when a patient is present in the interaction. Furthermore, since this study controlled for dialect, Black individuals speaking dialects other than GAE may also further elicit a clinician's racial bias and result in increased biases within the auditory-perceptual evaluations of voice.

Limitations and Future Study Directions

One limitation to this study was that all the samples were rated by the listeners in one block of time, which may have contributed to some familiarity when a speaker was played later in the ratings with a different labeled race. Although the prior work by Norotsky et al. (2023) completed ratings in multiple sessions to reduce this limitation, the feasibility of this design with 30 active SLPs seemed to present an unreasonable burden. To mitigate the risk of listeners identifying individual voices across conditions, voices were randomized within condition sets to reduce the likelihood of remembering orders of voice samples and, thus, identifying individual samples from these order strands.

Another limitation of this study was that many more of the speakers were females (n = 34) than males (n = 6). The rationale for this was to accurately represent the sex ratio of individuals presenting with voice disorders. However, this significantly higher proportion of female speakers eliminated the ability to detect any potential interaction between sex (or gender) and racial bias in auditory-perceptual evaluations. Future work should examine if such an interaction exists, since this would not be possible to extrapolate from the current methodology.

Auditory-perceptual evaluation of voice is one small piece of the scope of a voice-centric SLP's practice. There are other components within evaluation and intervention procedures that are susceptible to the impact of bias, such as the referral practices from otolaryngologists recommending voice therapy; establishing treatment choices and targets (e.g., specific therapy protocols and programs); and recommendations for the frequency, intensity, and duration of intervention. Future work should assess for minimizing biases for Black individuals across the complete span of voice care.

Conclusions

This study found that voice-focused SLPs demonstrated a small minimizing bias via auditory-perceptual evaluation of OS when a Black speaker's race was labeled as Black relative to when it was labeled as White. However, the clinical significance of this small reduction in OS scores may be minimal. The degree of bias did not relate to the difference in OS when speakers were labeled as Black or White. To mitigate the impact of implicit racial bias on auditory-perceptual impressions of dysphonia, we recommend discussions, reflection, and antiracism training to heighten awareness and improve culturally responsive care for Black patients seeking voice evaluation and therapy services.

Data Availability Statement

The data sets generated and/or analyzed during this study are not publicly available due to the inability to fully deidentify voice recordings but are available from the senior author on reasonable request.

Acknowledgments

This work was supported by National Institute on Deafness and Other Communication Disorders Grant R01 DC015570 (awarded to Principal Investigator: Cara E. Stepp).

References

Arora, S., Stouffer, G. A., Kucharska-Newton, A., Vaduganathan, M., Qamar, A., Matsushita, K., 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), Article e010203. https://doi.org/10.1161/JAHA.118.010203

Bergmark, R. W., & Sedaghat, A. R. (2017). Disparities in health in the United States: An overview of the social determinants of health for otolaryngologists. *Laryngoscope Investigative Otolaryngology, 2*(4), 187–193. https://doi.org/10.1002/lio2.81

Black, L. I., Vahratian, A., & Hoffman, H. J. (2015). *Communication disorders and use of intervention services among children aged 3–17 years: United States, 2012. NCHS data brief. Number 205*. Centers for Disease Control and Prevention.

Boone, D. R., McFarlane, S. C. L. V. B. S., & Zraick, R. I. (2014). *The voice and voice therapy*. Pearson.

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

Centola, D., Guilbeault, D., Sarkar, U., Khoong, E., & Zhang, J. (2021). The reduction of race and gender bias in clinical treatment recommendations using clinician peer networks in an experimental setting. *Nature Communications, 12*(1), Article 6585. https://doi.org/10.1038/s41467-021-26905-5

Charity, A. H. (2008). African American English: An overview. *Population, 15*(2), 33–42. https://doi.org/10.1044/cds15.2.33

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillside, NJ: Lawrence Erlbaum Associates.

Cooper, L. A., Saha, S., & Van Ryn, M. (2022). Mandated implicit bias training for health professionals—a step toward equity in health care. *Paper presented at the JAMA Health Forum, 3*(8), Article e223250. https://doi.org/10.1001/jamahealthforum.2022.3250

Damrose, J. F., Goldman, S. N., Groessl, E. J., & Orloff, L. A. (2004). The impact of long-term botulinum toxin injections on symptom severity in patients with spasmodic dysphonia. *Journal of Voice, 18*(3), 415–422. https://doi.org/10.1016/j.jvoice.2000.11.001

Davidson, M. M., Alonzo, C. N., & Stransky, M. L. (2022). Access to speech and language services and service providers for children with speech and language disorders. *American Journal of Speech-Language Pathology, 31*(4), 1702–1718. https://doi.org/10.1044/2022_AJSLP-21-00287

Duello, T. M., Rivedal, S., Wickland, C., & Weller, A. (2021). Race and genetics versus "race" in genetics: A systematic review of the use of African ancestry in genetic studies. *Evolution, Medicine, and Public Health, 9*(1), 232–245. https://doi.org/10.1093/emph/eoab018

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

Eadie, T. L., Kapsner, M., Rosenzweig, J., Waugh, P., Hillel, A., & Merati, A. (2010). The role of experience on judgments of dysphonia. *Journal of Voice, 24*(5), 564–573. https://doi.org/10.1016/j.jvoice.2008.12.005

Eadie, T. L., & Kapsner-Smith, M. (2010). The effect of listener experience and anchors on judgments of dysphonia. *Journal of Speech, Language, and Hearing Research, 54*(2), 430–447. https://doi.org/10.1044/1092-4388(2010/09-0205)

Eadie, T. L., Nicolici, C., Baylor, C., Almand, K., Waugh, P., & Maronian, N. (2007). Effect of experience on judgments of adductor spasmodic dysphonia. *The Annals of Otology, Rhinology, and Laryngology, 116*(9), 695–701. https://doi.org/10.1177/000348940711600912

Ellis, C. (2009). Does race/ethnicity really matter in adult neurogenics? *American Journal of Speech-Language Pathology, 18*(4), 310–314. https://doi.org/10.1044/1058-0360(2009/08-0039)

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 ed.). Harper & Row.

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

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, A. G., Banaji, M. R., & Nosek, B. A. (n.d.). *Race Implicit Association Test*. Project implicit. Retrieved April 7, 2024, from 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

Hagiwara, N., Kron, F. W., Scerbo, M. W., & Watson, G. S. (2020). A call for grounding implicit bias training in clinical and translational frameworks. *The Lancet, 395*(10234), 1457–1460. https://doi.org/10.1016/S0140-6736(20)30846-1

Hapner, E., Portone-Maira, C., & Johns, M. M., III. (2009). A study of voice therapy dropout. *Journal of Voice, 23*(3), 337–340. https://doi.org/10.1016/j.jvoice.2007.10.009

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/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/10.1016/j.jpain.2015.03.003

Hur, K., Zhou, S., Bertelsen, C., & Johns, M. M., III. (2018). Health disparities among adults with voice problems in the United States. *The Laryngoscope, 128*(4), 915–920. https://doi.org/10.1002/lary.26947

Javitt, J. C., McBean, A. M., Nicholson, G. A., Babish, J. D., Warren, J. L., & Krakauer, H. (1991). Undertreatment of glaucoma among Black Americans. *Ew 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,* Article 100563. https://doi.org/10.1016/j.ssmph.2020.100563

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press. https://doi.org/10.1017/CBO9780511809477

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)

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Laczi, E., Sussman, J. E., Stathopoulos, E. T., & Huber, J. (2005). Perceptual evaluation of hypernasality compared to HONC measures: The role of experience. *The Cleft Palate Craniofacial Journal, 42*(2), 202–211. https://doi.org/10.1597/03-011.1

Lugo, V. A., Wood, C., Torres-Chavarro, M., & Garcia-Salas, M. (2023). Barriers to graduate applications: Perspectives of undergraduate students from underrepresented backgrounds in communication science and disorders. *Perspectives of the ASHA Special Interest Groups, 8*(4), 729–740. https://doi.org/10.1044/2023_PERSP-22-00142

Mahendra, N., & Spicer, J. (2014). Access to speech-language pathology services for African American clients with aphasia: A qualitative study. *Division 14 Newsletter, 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 and Medicine, 199,* 219–229. https://doi.org/10.1016/j.socscimed.2017.05.009

Marcelin, J. R., Siraj, D. S., Victor, R., Kotadia, S., & Maldonado, Y. A. (2019). The impact of unconscious bias in healthcare: How to recognize and mitigate it. *The Journal of Infectious Diseases, 220*(Suppl. 2), S62–S73. https://doi.org/10.1093/infdis/jiz214

Mayne, G. V., & Namazi, M. (2023). Social determinants of health: Implications for voice disorders and their treatment. *American Journal of Speech-Language Pathology, 32*(3), 1050–1064. https://doi.org/10.1044/2023_AJSLP-21-00183

Meissner, F., Grigutsch, L. A., Koranyi, N., Müller, F., & Rothermund, K. (2019). Predicting behavior with implicit measures: Disillusioning findings, reasonable explanations, and sophisticated solutions. *Frontiers in Psychology, 10,* Article 2483. 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

Misono, S., Marmor, S., Roy, N., Mau, T., & Cohen, S. M. (2017). Factors influencing likelihood of voice therapy attendance: Report from the CHEER network. *Otolaryngology—Head and Neck Surgery, 156*(3), 518–524. https://doi.org/10.1177/0194599816679941

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

Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2012). Are minority children disproportionately represented in early intervention and early childhood special education? *Educational Researcher, 41*(9), 339–351. https://doi.org/10.3102/0013189X12459678

Morgan, P. L., Hammer, C. S., Farkas, G., Hillemeier, M. M., Maczuga, S., Cook, M., & Morano, S. (2016). Who receives speech/language services by 5 years of age in the United States? *American Journal of Speech-Language Pathology, 25*(2), 183–199. https://doi.org/10.1044/2015_AJSLP-14-0201

Morton, M. E., Gibson-Young, L., & Sandage, M. J. (2022). Framing disparities in access to medical speech-language pathology care in rural Alabama. *American Journal of Speech-Language Pathology, 31*(6), 2847–2860. https://doi.org/10.1044/2022_AJSLP-22-00025

Morton, M. E., & Sandage, M. J. (2022). Sex and race reporting and representation in noncancerous voice clinical trials: A meta-analysis of National Institutes of Health–registered research between 1988 and 2021. *Journal of Speech, Language, and Hearing Research, 65*(7), 2594–2607. https://doi.org/10.1044/2022_JSLHR-22-00141

Morton-Jones, M. E., & Timmons Sund, L. (2023). Exploring culturally responsive voice care: Mitigating disparities by addressing implicit bias. *Perspectives of the ASHA Special Interest Groups, 9*(5), 1283–1289. https://doi.org/10.1044/2023_PERSP-23-00137

National Academies of Sciences, Engineering, and Medicine. (2019). *Criteria for selecting the leading health indicators for healthy people 2030*. National Academies Press. https://doi.org/10.17226/25531

Nelson, A. (2002). Unequal treatment: Confronting racial and ethnic disparities in health care. *Journal of the National Medical Association, 94*(8), 666–668.

Nelson, M., & Wilson, L. (2021). Implicit bias and multilingual assessment in school-based speech-language pathologists. *Perspectives of the ASHA Special Interest Groups, 6*(6), 1690–1704. https://doi.org/10.1044/2021_PERSP-21-00113

Norotsky, R. L., Dahl, K. L., Cocroft, S., Sauder, C., Tracy, L. F., & Stepp, C. E. (2023). Does implicit racial bias affect auditory-perceptual evaluations of dysphonic voices? *Journal of Voice*. Advance online publication. https://doi.org/10.1016/j.jvoice.2023.11.023

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

Othman, L. B. (2018). The problem of disproportional representation of students from minority races in special education. *International Journal of Special Education, 33*(1), 171–183.

Papandrea, M. T., Namazi, M., Ghanim, I., & Patten, S. (2023). Identifying racial and socioeconomic biases in New Jersey special education eligibility. *Language, Speech, and Hearing Services in Schools, 54*(2), 600–617. https://doi.org/10.1044/2022_lshss-22-00138

Patel, R. R., Awan, S. N., Barkmeier-Kraemer, J., Courey, M., Deliyski, D., Eadie, T., Paul, D., Švec, J. G., & Hillman, R. (2018). Recommended protocols for instrumental assessment of voice: American Speech-Language-Hearing Association expert panel to develop a protocol for instrumental assessment of vocal function. *American Journal of Speech-Language Pathology, 27*(3), 887–905. https://doi.org/10.1044/2018_AJSLP-17-0009

Pope, L., Light, J., & Franklin, A. (2022). Black children with developmental disabilities receive less augmentative and alternative communication intervention than their White peers: Preliminary evidence of racial disparities from a secondary data analysis. *American Journal of Speech-Language Pathology, 31*(5), 2159–2174. https://doi.org/10.1044/2022_ajslp-22-00079

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' ratings of quality of life. *American Journal of Medicine, 108*(7), 561–566. https://doi.org/10.1016/s0002-9343(00)00352-1

Sapienza, C. M. (1997). Aerodynamic and acoustic characteristics of the adult African American voice. *Journal of Voice, 11*(4), 410–416. https://doi.org/10.1016/s0892-1997(97)80036-7

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 of Psychological Science, 16*(2), 435–442. https://doi.org/10.1177/1745691621991860

Sim, W., Lim, W. H., Ng, C. H., Chin, Y. H., Yaow, C. Y. L., Cheong, C. W. Z., Khoo, C. M., Samarasekera, D. D., Devi, M. K., & Chong, C. S. (2021). The perspectives of health professionals and patients on racism in healthcare: A qualitative systematic review. *PLOS ONE, 16*(8), Article e0255936. https://doi.org/10.1371/journal.pone.0255936

Smedley, A., & Smedley, B. D. (2005). Race as biology is fiction, racism as a social problem is real: Anthropological and historical perspectives on the social construction of race. *American Psychologist, 60*(1), 16–26. https://doi.org/10.1037/0003-066x.60.1.16

Stachler, R. J., Francis, D. O., Schwartz, S. R., Damask, C. C., Digoy, G. P., Krouse, H. J., McCoy, S. J., Ouellette, D. R., Patel, R. R., Reavis, C. W., Smith, L. J., Smith, M., Strode, S. W., Woo, P., & Nnacheta, L. C. (2018). Clinical practice guideline: Hoarseness (dysphonia) (update). *Head and Neck Surgery, 158*(Suppl. 1), S1–S42. https://doi.org/10.1177/0194599817751030

Sun, M., Oliwa, T., Peek, M. E., & Tung, E. L. (2022). Negative patient descriptors: Documenting racial bias in the electronic health record. *Health Affairs, 41*(2), 203–211. https://doi.org/10.1377/hlthaff.2021.01423

Van Ryn, M., & Burke, J. (2000). The effect of patient race and socio-economic status on physicians' perceptions of patients. *Social Science & Medicine, 50*(6), 813–828. https://doi.org/10.1016/s0277-9536(99)00338-x

Vora, S., Dahlen, B., Adler, M., Kessler, D. O., Jones, V. F., Kimble, S., & Calhoun, A. (2021). Recommendations and guidelines for the use of simulation to address structural racism and implicit bias. *Simulation in Healthcare, 16*(4), 275–284. https://doi.org/10.1097/sih.0000000000000591

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

Yu, B., Horton, R., Munson, B., Newkirk-Turner, B. L., Johnson, V. E., Khamis-Dakwar, R., Muñoz, M. L., & Hyter, Y. D. (2022). Making race visible in the speech, language, and hearing sciences: A critical discourse analysis. *American Journal of Speech-Language Pathology, 31*(2), 578–600. https://doi.org/10.1044/2021_AJSLP-20-00384