Listener Age and Gender Diversity: Effects on Voice-based Perception of Gender

*Katherine M. Brown, *Kimberly L. Dahl, †Gabriel J. Cler, and *,‡,§Cara E. Stepp, *‡§Boston, Massachusetts, and †Oxford, UK

Summary: Objective. An important clinical outcome of voice masculinization treatments in transmasculine speakers is voice-based perception of gender. Rigorous assessments of voice treatment that utilize ratings of perception of gender typically do not control for demographic characteristics of the listeners. The objective of the present study was to determine the effect of listeners' age and gender diversity on voice-based judgments of speaker gender.

Methods. Speech stimuli were produced by a single transmasculine individual over approximately one year of hormone replacement therapy, during which he experienced significant changes in his voice. Three groups of listeners rated speech stimuli on a visual analog scale with anchors ranging from "definitely male" to "guessing male" to "guessing female" to "definitely female." Listener groups were N = 10 cisgender young adults, N = 10 cisgender older adults, and N = 10 gender diverse individuals.

Results. All groups rated the speaker as consistently female through week 14 of hormone replacement therapy and consistently male after week 28. Mean responses of the three groups of listeners were highly correlated (Pearson's correlations all $r > 0.97$).

Conclusion. Given reasonable group sizes, average ratings of gender perception of a transmasculine speaker are not highly influenced by varying listener age and gender minority status.

Key Words: Gender Minority−Perception of Gender−Voice−Transgender−Age.

INTRODUCTION

Congruence of voice and gender identity is a significant concern for many transmasculine individuals.[1] It is often assumed that this congruence will be achieved via the voice masculinizing effects of hormone replacement therapy (HRT) with exogenous testosterone.[2,3] Yet there is a relative scarcity of evidence on the effectiveness of HRT for achieving this outcome. In fact, as many as 31% of transmasculine individuals who have initiated HRT remain dissatisfied with their voice masculinity.[1] Much of the existing literature has sought to document HRT-induced voice masculinization by measuring changes in voice acoustics, such as fundamental frequency (f_o),[4-9] phonatory frequency range,[4,7,8,10] and measures of voice quality (e.g., noise-to-harmonics ratio,[4,10] jitter,[8,10] and shimmer[4,8,10]). An alternative outcome measure, employed by some studies[7,8,10,11] is the self-perception of voice of the speaker.

Listener perception of gender is also important to assess, as being perceived by others as one's true gender is critical to the safety and overall health of transgender individuals.[12-14] Research on transgender voice thus commonly includes listener perceptual tasks as a standard component.[11,15-23] The listeners completing these perceptual evaluations of voice are often young, cisgender adults.[17,21-24] This is likely a result of convenience samples in which participant recruitment is largely targeted at undergraduate students. Older adults are thus less likely to receive notice of study participation opportunities. They are also more likely to fail to meet inclusionary criteria due to the greater prevalence of hearing loss among older adults.[25,35] Gender minorities may also be underrepresented among study participants due to a combination of their constituting a smaller proportion of the general population and a lack of targeted recruitment efforts by researchers. This practice of convenience sampling thus precludes a more diverse set of listeners that would include older adults and gender minorities. The impact of this methodological approach on research outcomes is unknown. If listener characteristics such as age or gender were to affect the listener's perception of others, the conclusions drawn from research based solely on the perceptions of a narrow segment of the population would not be generalizable.

Some related research suggests that potential age effects warrant further investigation. In one study,[25] older adults showed reduced accuracy in recalling the gender of a speaker. These results were interpreted as suggestive of the impact of age-related cognitive changes on the integration of nonsemantic content, like gendered characteristics of voice. In another study,[26] older adults and younger adults performed a listening task in which both groups were presented with voice samples from cisgender individuals of varying ages; the listeners were asked to determine the speakers' genders and make judgments with regards to speakers' voice quality such as naturalness, clarity, and ease of understanding. While both groups identified the speakers' genders correctly the majority of the time, young adults were more likely to misinterpret older cisgender women's voices to be that of younger cisgender men's voices. The

Accepted for publication February 6, 2020.

From the *Department of Speech, Language, and Hearing Sciences, Boston University, Boston, Massachusetts; †Department of Experimental Psychology, University of Oxford, Oxford, UK; ‡Department of Otolaryngology − Head and Neck Surgery, Boston University School of Medicine, Boston, Massachusetts; and the §Department of Biomedical Engineering, Boston University, Boston, Massachusetts.

Address correspondence and reprint requests to Cara E. Stepp, 635 Commonwealth Avenue, Boston, MA 02215. E-mail: cstepp@bu.edu

Journal of Voice, Vol. ■■, No. ■■, pp. ■■−■■
0892-1997

© 2020 The Voice Foundation. Published by Elsevier Inc. All rights reserved.

https://doi.org/10.1016/j.jvoice.2020.02.004

older adult group made less errors in identifying speakers' genders. The varying error rates in this gender identification task between the two groups support that listener characteristics, such as age, may impact perceptual ratings of voice-based perception of gender.

Few studies have specifically investigated whether listener characteristics of gender identity or sexual orientation might extend beyond general attitudes to voice-based perceptions of gender-diverse individuals. In one study,[27] investigators found a small effect of sexual orientation on listeners' femininity ratings of transfeminine individuals such that the nonstraight listeners rated the transfeminine speakers' voice recordings as significantly more feminine than did the straight listeners.[27] The authors of this study proposed that nonstraight listeners may have greater awareness of their own voices, thereby translating into a greater sensitivity in auditory perception than that of straight listeners. Another study[28] found no effect of LGBTQ status on gender perceptions. Others have demonstrated an effect of gender on the accuracy of listeners' gender perceptions.[29] It is thus unclear whether individual characteristics, such as sexual orientation or gender identity, may influence perceptions of gender.

The goal of the present study was to determine the impact of listeners' age and gender diversity on voice-based perceptions of speaker gender. Based on limited prior research in this domain, this study was exploratory in nature, investigating the potential role that age and gender diversity may or may not have in the perception of gender of a transmasculine speaker. Speech stimuli were produced by a single transmasculine individual over approximately 1 year of HRT, during which he experienced significant changes in his voice. These changes are fully described in our previous work[30] and included decreased mean f_o, downward shift in f_o range, decreased fourth formant frequency, and other changes in voice quality, physiology, and perceptual characteristics. Three groups of listeners were incorporated: a group of young cisgender adults, a group of older cisgender adults, and a group of individuals who are gender diverse.

MATERIAL AND METHODS

Speech stimuli

The speech stimuli in this study were provided by a single transmasculine individual (assigned female at birth) over approximately 1 year of HRT. Speech samples were recorded three times in the 2 weeks prior to initiation of HRT and every 2 weeks thereafter for 1 year. The participant was a native American English speaker and reported no speech, language, or hearing disorder, nor any neurological or cognitive disorders. He did not undergo any voice therapy during the course of data collection. Detailed descriptions of the speaker and the acoustic measures derived from the speech samples are reported elsewhere.[23]

The speech recordings were collected in a sound-attenuating audiometric booth. The speaker wore a head-mounted microphone (Shure Omni WH20) that rested 6-10 cm from his mouth at a 45° angle from the midline, as recommended by the American Speech-Language-Hearing Association (ASHA).[31] Acoustic signals were preamplified via an RME Quadmic II (RME, Haimhausen, Germany) and sampled at 44,100 Hz with 16-bit resolution using a MOTU UltraLite-mk3 Hybrid (model UltraLite3Hy; MOTU, Cambridge, MA) and SONAR software (Cakewalk, Boston, MA). The stimuli used for this study were acoustic recordings of excerpts of the Rainbow Passage[32] (sentences 2–4). The duration of each stimulus was an average of 13.2 seconds (range 12.6–14.6 seconds). Stimuli were normalized for peak intensity using MATLAB.

Listeners

Thirty listeners participated in the study and were divided into three listener groups: young cisgender adults (YCAs), older cisgender adults (OCAs), and gender-diverse adults (GDAs). The YCAs in this study were distinct from the young adult raters in our previous work.[23] All listeners were native American English speakers who reported no neurological or cognitive disorders and no current speech, language, or hearing disorders. One listener reported a diagnosis of laryngeal hyperfunction, and one listener reported a childhood speech delay. Cisgender adults were recruited via Boston University job board and flyers. GDAs were recruited via flyers at local LGBTQ events and internet postings on LGBTQ social media groups. All listeners completed informed consent in compliance with the Boston University Institutional Review Board.

Listeners were YCAs (five cisgender females, five cisgender males, M: 21.8 years, SD: 2.3, range: 19–27), OCAs (five cisgender females, five cisgender males, M: 63.5 years, SD: 5.6, range: 54–72), and GDAs (five transgender males, five nonbinary individuals, M: 31 years, SD: 9.3, range: 20–53). Nine GDA participants were assigned female at birth and one was assigned male at birth. Five GDAs self-identified as transmasculine, two as nonbinary, and three as nonbinary transmasculine. Pronouns of the GDAs were "he/him/his" (five participants) and "they/them/theirs" (five participants).

The sexual orientation of the three groups were as follows. In the YCA group, nine listeners reported their sexual orientation as either "straight" or "heterosexual" and one cisgender female listener did not report. In the OCA group, nine listeners reported their sexual orientation as "straight" or "heterosexual" and one cisgender male listener reported "bi." In the GDA group, three transmasculine listeners reported their sexual orientation as "gay." Three listeners reported their sexual orientation as "queer," who were transmasculine, nonbinary transmasculine, and nonbinary. One transmasculine individual reported "homoflexible" and one nonbinary transmasculine listener reported "gray-ace panromantic." One nonbinary person reported "bisexual" and one nonbinary person reported "lesbian."

Hearing and cognitive screening

All participants underwent a hearing screening in which they were presented pulsed pure tones at octaves from 125

to 8000 Hz in both ears.[33,34] All YCAs passed the hearing screening at 25 dB at these frequencies with the exception of one YCA who did not hear 8000 Hz at 50 dB in the right ear only. Three OCAs heard all tones at 25 dB in both ears. Five OCAs heard frequencies from 125 to 4000 Hz at a minimum of 30 dB or better in both ears. Two OCAs heard 125 Hz to 4000 Hz in at least one ear at 40 dB or better. Eight GDAs heard all tones at 25 dB in both ears. Two GDAs heard all tones at 25 dB in at least one ear; they detected tones from 125 to 4000 Hz in both ears at least 30 dB HL or better. Participants in all groups passed a minimum threshold of 40 dB HL from 125 to 4000 Hz in at least one ear and 50 dB at 8000 Hz in at least one ear. These hearing thresholds were selected to accommodate a representative sample of older adults given the prevalence of hearing loss in this population.[35,36]

Individuals over 50 years old completed the Montreal Cognitive Assessment (MoCA)[37] immediately after consent. An inclusionary MoCA score of ≥22 was set to ensure a level of cognitive functioning sufficient for informed consent. All participants met this criterion.

Perceptual listening task

All study procedures took place in a sound-attenuating audiometric booth produced by IAC Acoustics. Participants sat in front of a computer monitor and wore circumaural Sennheiser 280 Pro headphones (Sennheiser Electronic Corporation). Listeners were played two example recordings (one from a cisgender male speaker and one from a cisgender female speaker) prior to the experimental task for the purpose of setting the volume to a comfortable level. The order of these sound files was counterbalanced across participants.

Participants were told that the present study sought to examine which voice features may contribute to the assumptions a listener makes about a speaker's identity. The participants then listened to the speech samples described above and were instructed to indicate the speaker's gender and their confidence in that decision by placing a marker on a 100 mm visual analog scale (VAS). The VAS had anchors ranging from "definitely male" to "definitely female" with intermediate anchors as shown in Figure 1. The 26 recordings were randomized and played twice; each participant thus rated 52 recordings. Listeners could not leave the marker at the midline, but could otherwise place it anywhere along the VAS. This listening task took approximately 20 minutes to complete.

Analysis

Intrarater reliability was assessed for each listener by calculating Pearson's correlation coefficients for ratings of repeated stimuli. No listeners were excluded based on reliability given the variable nature of gender perception.[3] Interrater reliability was calculated for each group using intraclass correlation coefficient (C,k). Between-group differences were assessed with Pearson's correlation coefficients between group means for all timepoints.

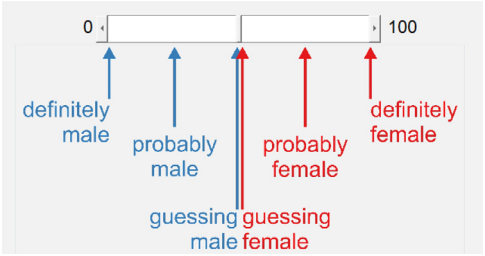

FIGURE 1. The custom-designed visual analog scale presented to listeners for the perceptual study. "Definitely male" was at 0 mm, "probably male" at 27 mm, "guessing male" at 49 mm, "guessing female" at 51 mm, "probably female" at 73 mm, and "definitely female" at 100 mm.

RESULTS

Intrarater reliability ranged from 0.42 to 0.97 (see Table 1). Across groups, the mean intrarater reliability was similar (YCA: 0.87; OCA: 0.76; GDA: 0.80), but the YCA group had a narrower range of reliability (YCA minimum: 0.70; OCA minimum: 0.43; GDA minimum: 0.42). Interrater reliability was strong for all groups (YCA: 0.97; OCA: 0.95; GDA: 0.98).

Mean gender ratings for all groups are shown in Figure 2, and individual gender ratings for the three groups are shown in Figure 3. Visual inspection of the mean ratings shows that the error bars overlap for the entire time course of HRT (x-axis). Statistical analysis indicates a strong concordance between all groups: $r_{YCA-OCA} = 0.988$; $r_{YCA-GDA} = 0.985$; $r_{OCA-GDA} = 0.976$.

DISCUSSION

The present study aimed to investigate the potential effect that listener age and gender minority status may or may not have on individuals' perception of gender. By comparing YCA, OCA, and GDA ratings of gender of one transmasculine speaker's voice over the course of HRT, this study aimed to tease apart whether age and gender minority status play a role in gender perception of transmasculine voice. If differences in listener perception were found due to age or gender minority status, this would have important implications for research examining transmasculine voice, which may rely on ratings such as these to evaluate treatment efficacy.

Mean ratings of gender are similar, regardless of group

All groups rated the speaker as consistently female through week 14 of HRT and consistently male after week 28. These mean responses of the three groups of listeners are depicted in Figure 2. OCA mean ratings compared to YCA mean ratings were highly correlated, supporting the conclusion that age differences do not yield different average responses. Average GDA ratings compared to OCA and YCA responses were also highly correlated; this suggests that gender minority status may not impact mean ratings of voice-based gender perception.

TABLE 1.
Intrarater and interrater reliability for the three groups

	Young cisgender adults	Older cisgender adults	Gender diverse adults
Intrarater reliability	Range: 0.70 – 0.96 (M = 0.87; SD = 0.08)	Range: 0.43 – 0.97 (M = 0.76; SD = 0.16)	Range: 0.42 – 0.97 (M = 0.80; SD = 0.16)
Interrater reliability	0.97	0.95	0.98

Overall, these results suggest that, given a minimum group size of 30 listeners and satisfactory hearing amongst participants, average ratings of voice-based gender perception of a transmasculine speaker may not be highly susceptible to influence of listener age and gender minority status. The results of the present study suggest that one may reasonably assume it to be unnecessary to recruit a particular age and gender-diverse listener composition in order to gather statistically valid average responses for use in research on transmasculine voice perception.

Notably, mean responses were the same across the three groups regardless of gender minority status and age. However, recruitment of the OCA group was particularly problematic. Prior to participation in the present study, all potential participants underwent a prescreening questionnaire over the phone or online to determine eligibility. Several OCAs were excluded in the prescreening process due to reporting known hearing issues. At the study site, three OCAs were excluded due to hearing status: two failed the hearing screening and did not participate in the perceptual listening task, and another reported tinnitus after the hearing screening but did not note this on the prescreening questionnaire. Moreover, of the ten OCAs included in the present study, only three heard all frequencies from 125 Hz to 8000 Hz at 25 dB HL. The remaining seven OCAs included in the study heard 125 Hz to 4000 Hz in at least one ear at 45 dB HL or better, and 8000 Hz at 50 dB HL in at least one ear or better. Although the study was not powered to compute statistical contrasts for age, gender, sexual orientation, and hearing status, visual inspection of individual data did not reveal an obvious relationship between OCA gender, hearing status, sexual orientation, and perception of gender. Upon visual inspection of the data, the ratings of the one nonstraight individual in the OCA group did not differ meaningfully from the ratings of others in the OCA group, and similarly, the ratings of the individual in the YCA group whose sexual orientation was not reported did not differ meaningfully from others in the YCA group. Future research that includes a larger sample size representative of the entire range of ages and gender identities would allow for further exploration of this question.

One potential explanation for the lack of differences in mean ratings across the three groups is that gender perception may be based upon stereotypes about vocal gender that are shared by individuals in a particular society regardless of age and gender identity. The listeners in the present study were all native speakers of American English and resided in the greater Boston area, and thus may have been subjected to similar learned stereotypes and prejudices about vocal gender. This listening task may have activated the participants' automatic stereotypes[38] about "definitely male" voices and "definitely female"

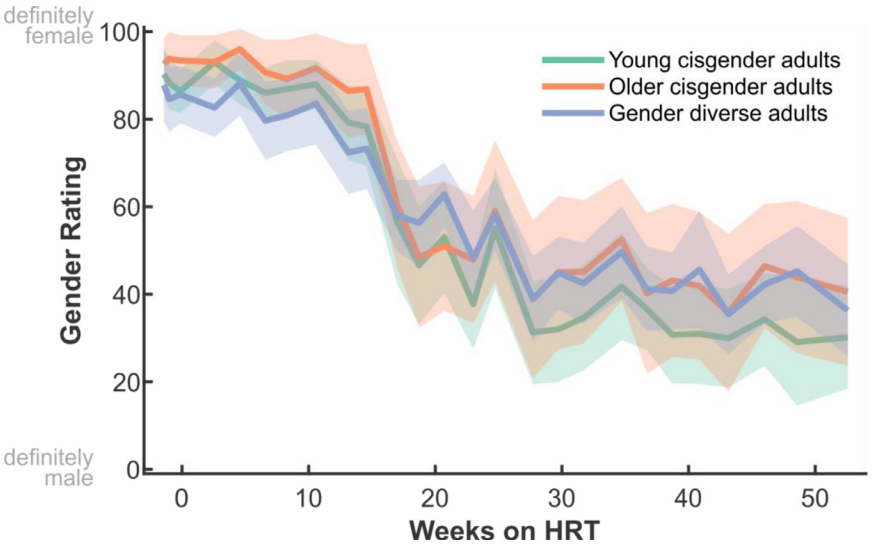

FIGURE 2. Mean gender ratings across groups. Young cisgender listeners are shown in green; older cisgender listeners are shown in orange, and gender diverse listeners are shown in purple. Shaded regions show standard error for each group.

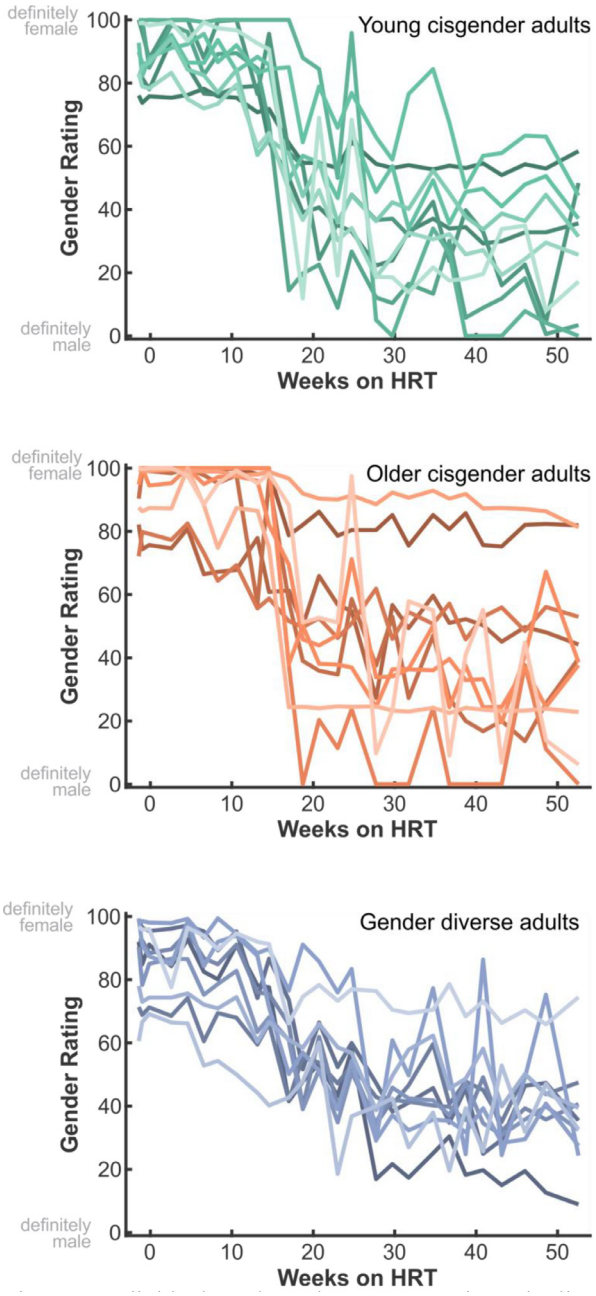

FIGURE 3. Individual gender ratings. Young cisgender listeners are shown on top plot (green); older cisgender listeners are shown in middle plot (orange), and gender diverse listeners are shown in bottom plot (purple). Different shades represent individual listeners.

voices, explaining why no differences were found due to listener age or gender identity in this study.

It is worth noting that the inclusion of data from listeners with relatively low intrarater reliability was a deliberate choice. While low intrarater reliability may warrant elimination of the unreliable rater's data in some applications, this may not be justified in the case of voice-based gender perception. Following the argument of Azul and colleagues,[3] the act of gendering measured in this study was considered an inherently variable and complex process in which both the speaker and listener play active roles. In this view, reliability measures are not meant to reflect accuracy across responses, but rather may hint at individual differences in how rigidly a listener adheres to expectations of gender expression.

Effects of age on gender perception

This study investigated whether YCAs compared to OCAs would rate the gender of one transmasculine speaker differently; differences in listener perception due to age would have important implications for research examining transmasculine voice, which may rely on ratings such as these to evaluate treatment efficacy. Despite studies pointing to potential similarity (in the US)[39] and dissimilarity (in Hong Kong[40] and Sweden[41]) of attitudes harbored by younger and older adults towards the LGBTQ community, the mean gender perceptions of these two groups in the present study were the same.

Concerns about age composition of listeners in gender perceptual studies may be rooted in stereotypes surrounding age and conservativism.[42] While age may be a predictor of conservatism in general[42] and attitudes toward gender minorities specifically,[40,41] our results do not suggest that such effects carry over into voice-based gender perception. Rather, if conservatism or attitudes toward gender minorities do have any effect on gender perception, it would not be captured by simply recruiting older listeners.

Impact of gender diversity on gender perception

We found that GDAs did not rate the gender of a transmasculine speaker differently than cisgender listeners. Our results are in line with previous work examining gender perception among gender and sexual minority groups. The study conducted by Hannah[28] found that voice-based perception of gender by members of the LGBTQ community did not differ based on listener characteristics such as their sexuality and gender. This may be because this study was aimed at gathering the automatic assumptions made about the speakers' gender,[38] which are shared by individuals across a culture, and not aimed at gauging individuals explicit beliefs about transgender persons. The present study similarly does not support the assumption that gender perception by gender-diverse individuals differ in meaningful ways from those of cisgender individuals.

Limitations and future directions

This exploratory study involved only 30 listeners rating one transmasculine speaker's voice. Since the voice recordings used in this experiment were one single speaker, the findings cannot be generalizable to all transmasculine voices. Likewise, there may be other cultural factors that could influence perception of gender that were not investigated in the present study. Religiosity and education have been shown to impact attitudes toward transgender persons.[40] Additionally, the race/ethnicity of the transgender speaker has been shown to have an impact on voice-based gender classification.[43] The present study did not examine these factors. All the

participants were local to the greater Boston area, an urban setting, and were willing to come to the study site at a university to participate in a research study. In addition to potential lack of diversity in education, religiosity, and race, the urban environment may further limit generalizability of findings. For instance, LGBTQ persons in rural settings report more oppression and discrimination than those in urban settings.[44] This could indicate that the climate of urban settings may be more accepting towards LGBTQ persons, and thus the ratings of YCAs and OCAs in the present study may differ from those collected in a rural setting.

It is important to note that the gendering process investigated here matches an everyday event such as a phone call, in which voice serves as the primary cue for gender identity. This study does not capture exchanges in which visual information plays an important role in gender perception. In fact, such visual information has been shown to affect listener perception of gender.[20,45] These results are thus applicable to a specific type of interaction, one in which the voice stands apart from other aspects of gender expression. Future research that tracks changes in perception of gender using both auditory and visual stimuli is warranted.

CONCLUSIONS

The goal of the current study was to determine the effect of listeners' age and gender diversity on voice-based judgments of speaker gender. Three groups of listeners (YCAs, OCAs, and GDAs) rated speech stimuli from a transmasculine individual undergoing HRT from "definitely male" to "guessing male" to "guessing female" to "definitely female." All groups rated the speaker as consistently female through week 14 of HRT and consistently male after week 28. Although individual responses of listeners varied in all three groups, mean responses of the three groups of listeners were highly correlated (Pearson's correlations all $r > 0.97$). Thus, these results suggest that, given reasonable group sizes, average ratings of gender of transmasculine speakers are not highly influenced by varying age and gender minority status. Therefore, given the infeasible time and expense in recruiting large groups of older adults who do not have hearing loss to provide gender ratings, it may not be necessary to control for listeners' age and gender diversity in future studies in order to achieve valid estimates of gender perception.

Acknowledgments

This work was supported by the grants DC017637 (G.J.C.), DC014872 (G.J.C.), and DC015570 (C.E.S.) from the National Institute on Deafness and Other Communication Disorders (NIDCD). Thanks to Dante Cilento and Monique Tardif for assistance with data collection and Manuel Diaz Cadiz for assistance with data processing.

REFERENCES

1. Van Borsel J, De Cuypere G, Rubens R, et al. Voice problems in female-to-male transsexuals. *Int J Lang Commun Disord*. 2000;35:427–442.
2. McNeill EJM, Wilson JA, Clark S, et al. Perception of voice in the transgender client. *J Voice*. 2008;22:727–733.
3. Azul D, Arnold A, Neuschaefer-Rube C. Do transmasculine speakers present with gender-related voice problems? *J Speech Lang Hear Res*. 2018;61:25–39.
4. Damrose EJ. Quantifying the impact of androgen therapy on the female larynx. *Auris Nasus Larynx*. 2009;36:110–112.
5. Deuster D, Di Vincenzo K, Szukaj M, et al. Change of speech fundamental frequency explains the satisfaction with voice in response to testosterone therapy in female-to-male gender dysphoric individuals. *Eur Arch Otorhinolaryngol*. 2016;273:2127–2131.
6. Irwig MS, Childs K, Hancock AB. Effects of testosterone on the transgender male voice. *Andrology*. 2017;5:107–112.
7. Nygren U, Nordenskjöld A, Arver S, et al. Effects of voice fundamental frequency and satisfaction with voice in trans men during testosterone treatment: a longitudinal study. *J Voice*. 2016;30. 766.e723-766.e734.
8. Van Borsel J, de Cuypere G, Rubens R, et al. Voice problems in female-to-male transsexuals. *Int J Lang Commun Disord*. 2000;35:427–442.
9. Zimman L. *Voices in Transition: Testosterone, Transmasculinity, and the Gendered Voice Among Female-to-Male Transgender People*. University of Colorado at Boulder; 2012. [Unpublished dissertation].
10. Hancock AB, Childs KD, Irwig MS. Trans male voice in the first year of testosterone: make no assumptions. *J Speech Lang Hear Res*. 2017;60:2472–2482.
11. Scheidt D, Kob M, Willmes K, et al. Do we need voice therapy for female-to-male transgenders? 2004.
12. Coleman E, Bockting W, Botzer M, et al. Standards of care for the health of transsexual, transgender, and gender-nonconforming people, version 7. *Int J Transgenderism*. 2012;13:165–232.
13. Hancock AB, Krissinger J, Owen K. Voice perceptions and quality of life of transgender people. *J Voice*. 2010;25:553–558.
14. Hays SE. *Attitudes About Voice and Voice Therapy Among Transgender Individuals*. Louisiana State University; 2013. [Unpublished master's thesis].
15. Gelfer MP, Schofield KJ. Comparison of acoustic and perceptual measures of voice in male-to-female transsexuals perceived as female versus those perceived as male. *J Voice*. 2000;14:22–33.
16. Hardy TLD, Rieger JM, Wells K, et al. Acoustic predictors of gender attribution, masculinity-femininity, and vocal naturalness ratings amongst transgender and cisgender speakers. *J Voice*. 2018.
17. Holmberg EB, Oates J, Dacakis G, et al. Phonetograms, aerodynamic measures, self-evaluations, and auditory perceptual ratings of male-to-female transsexual voice. *J Voice*. 2010;24:511–522.
18. Owen K, Hancock AB. The role of self- and listener perceptions of femininity in voice therapy. *Int J Transgenderism*. 2010;12:272–284.
19. Wolfe VI, Ratusnik DL, Smith FH, et al. Intonation and fundamental frequency in male-to-female transsexuals. *J Speech Hear Disord*. 1990;55:43–50.
20. Van Borsel J, de Pot K, De Cuypere G. Voice and physical appearances in female-to-male transsexuals. *J Voice*. 2009;23:494–497.
21. Hancock A, Colton L, Douglas F. Intonation and gender perception: applications for transgender speakers. *J Voice*. 2014;28:203–209.
22. Dahl KL, Mahler LA. Acoustic features of transfeminine voices and perceptions of voice femininity. *J Voice*. 2019.
23. Cler GJ, McKenna VS, Dahl KL, et al. Longitudinal case study of transgender voice changes under testosterone hormone therapy. [in press] *J Voice*. 2019. S0892199718305411.
24. King RS, Brown GR, McCrea CR. Voice parameters that result in identification or misidentification of biological gender in male-to-female transgender veterams. *Int J Transgenderism*. 2012;13:117–130.
25. Kausler DH, Puckett JM. Adult age differences in memory for sex of voice. *J Gerontol*. 1981;36:44–50.
26. Goy H, Pichora-Fuller KM, van Lieshout P. Effects of age on speech and voice quality ratings. *J Acoust Soc Am*. 2016;139:1648–1659.

ARTICLE IN PRESS

27. Hancock AB, Pool SF. Influence of listener characteristics on perceptions of sex and gender. *J Lang Soc Psychol*. 2017;36:599–610.
28. Hannah A. *Perception of Voice Gender by Members of the LGBTQ+ Community*. Speech, Language, and Hearing Sciences, University of Colorado; 2019. [Unpublished master's thesis].
29. Junger J, Habel U, Bröhr S, et al. More than just two sexes: The neural correlate of voice gender perception in gender dysphoria. *PLOS ONE*. 2014;9: e111672.
30. Cler GJ, McKenna VS, Dahl KL, et al. Longitudinal case study of transgender voice changes under testosterone hormone therapy. *J Voice*. 2019.
31. Patel Rita R, Awan Shaheen N, Barkmeier-Kraemer J, et al. Recommended protocols for instrumental assessment of voice: American speech-language-hearing association expert panel to develop a protocol for instrumental assessment of voice function. *Am J Speech-Language Pathol*. 2018;27:887–905.
32. Fairbanks G. *Voice and Articulation Drillbook*. 2nd ed. New York, NY: Harper & Row; 1960.
33. Association AS-L-H. Guidelines for manual pure-tone threshold audiometry. 2005.
34. Burk MH, Wiley TL. Continuous versus pulsed tones in audiometry. *Am J Audiol*. 2004;13:54061.
35. Cruickshanks KJ, Wiley TL, Tweed TS, et al. Prevalence of hearing loss in older adults in beaver dam, wisconsin. The epidemiology of hearing loss study. *Am J Epidemiol*. 1998;148:879–886.
36. Lee J, Dhar S, Abel R, et al. Behavioral hearing thresholds between 0.125 and 20 kHz using depth-compensated ear simulator calibration. *Ear Hear*. 2012;33:315–329.
37. Nasreddine ZS, Phillips NA, Bédirian V, et al. The montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005;53:695–699.
38. Devine PG. Stereotypes and prejudice: their automatic and controlled components. *J Pers Soc Psychol*. 1989;56:5.
39. Harvey J. *Attitudes of the General Population Toward Transsexuals*. Southern California University for Professional Studies; 2002.
40. King ME, Winter S, Webster B. Contact reduces transprejudice: a study on attitudes towards transgenderism and transgender civil rights in Hong Kong. *Int J Sex Health*. 2009;21:17–34.
41. Landén M, Innala S. Attitudes toward transsexualism in a Swedish national survey. *Arch Sex Behav*. 2000;29:375–388.
42. Grant MJ, Ross AS, Button CM, et al. Attitudes and stereotypes about attitudes across the lifespan. *Soc Behav Pers*. 2001;29:749–762.
43. Lagos D. Hearing gender: voice-based gender classification and transgender health inequality. *American Sociological Review*. 2019;84: 801–827.
44. Swank E, Fahs B, Frost DM. Region, social identities, and disclosure practices as predictors of heterosexist discrimination against sexual minorities in the United States. *Sociological Inquiry*. 2013;83:238–258.
45. Van Borsel J, De Cuypere G, Van den Berghe H. Physical appearance and voice in male-to-female transsexuals. *J Voice*. 2001;15:570–575.