

# The use of self-report measures to examine changes in perception in response to fittings using different signal processing parameters

Melinda Anderson<sup>1</sup>, Varsha Rallapalli<sup>2</sup>, Tim Schoof<sup>3</sup>, Pamela Souza<sup>2,4</sup>, and Kathryn Arehart<sup>5</sup>

<sup>1</sup>Department of Otolaryngology, University of Colorado Anschutz Medical Campus

<sup>2</sup>Department of Communication Sciences and Disorders, Northwestern University

<sup>3</sup>Department of Speech, Hearing and Phonetic Sciences, University College London

<sup>4</sup>Knowles Hearing Center, Northwestern University

<sup>5</sup>Department of Speech Language Hearing Sciences, University of Colorado at Boulder

## **Introduction**

Clinical best-practice guidelines recommend measuring hearing aid outcomes using standardized methods [1], including subjective outcomes via self-report measures. The usefulness of self-report measures for documenting changes in perceived performance *from unaided to aided listening* is well-established [2-8]. However, it is less clear whether and which measures might be sensitive to more subtle differences in perceived performance, such as when hearing aid signal processing features are modified [9-10]. Those modifications include wide dynamic range compression (WDRC) speed, strength and type of digital noise reduction, and frequency- lowering. Parameters for these features can vary greatly, even for the same audiogram. While audiologists tend to rely on the manufacturers' default fitting parameters [11], the ability to more systematically assess benefit -- including patients' self-perceived benefit -- could be used to more clearly direct programming adjustments. As such, it is of interest to explore whether self-report measures that query specific situations likely to be affected by

different signal processing are sensitive to differences in hearing aid fittings that vary in the way signal processing parameters are adjusted.

To this end, we retrospectively review a dataset collected as part of clinical trial [12; clinicaltrials.gov Identifier: NCT02448706]. Souza et al. [12] reported the results of a double-blinded randomized cross-over clinical trial in which all participants received two fittings with the same hearing aid. The first hearing aid fitting used mild processing (WDRC with slow time constants only). The second hearing aid fitting used strong processing (WDRC with fast time constants and with frequency compression). After wearing each fitting for at least four weeks participants filled out two different self-report measures. The first measure was the multidimensional Speech, Spatial, and Qualities of Hearing Scale (SSQ) [13], which was developed to assess hearing in complex environments and the components of spatial hearing. The second measure, the Effectiveness of Auditory Rehabilitation (EAR) [5] scale was designed to capture factors associated with hearing-related function and hearing-aid function. The EAR scale is composed of two modules, the Inner EAR and the Outer EAR. The Inner EAR is a pre- and post-fit measure that examines perception regarding hearing-related function. The Outer EAR is a post-fitting measure only, and captures factors related to hearing aids. For example, the measure includes questions about specific noises that may be affected by signal processing, such as wind noise and impact noise. The use of the two modules together provide a mechanism for documenting a more holistic view of the listener experience with regard to a specific hearing intervention. Our primary research aim was to determine whether self-perceived performance was sensitive to differences in signal processing. Our hypothesis was that participants would self-rate their perception

differently depending on the specific signal processing parameter settings associated with each fitting. Additionally, we hypothesized that these measures may quantify different facets of perceived benefit.

## **Methods**

Data were collected at two sites (University of Colorado Boulder and Northwestern University) as part of a registered clinical trial (clinicaltrials.gov Identifier: NCT02448706), using the same methods and equipment.

### **Participants**

Older adults with hearing loss were recruited for this study. In total, 49 older adults were consented. Of the consented 49 participants, 40 were included in final data analyses, aged 54 – 90 years (mean age 72 years, 19 women). Two participants dropped from the study before they were fit with hearing aids. Five participants dropped from the study shortly after their first hearing aid fitting (four due to an inability to acclimate to the hearing aid processing and one due to an inability to insert the aid). An additional two participants had to be excluded because they did not wear the hearing aids for at least four hours a day.

[figure 1 near here]

Participants all had sensorineural hearing loss (see Fig. 1), with air-bone gaps < 15 dB at octave frequencies from 0.5 to 4 kHz and normal tympanograms. Inclusion criteria were a high-frequency pure-tone average (HFPTA; 0.5, 1, 2, 4 kHz) of at least 30 dB HL, with audiometric thresholds up to 3 kHz  $\leq$  70 dB HL. Hearing losses were symmetrical, with the difference in HFPTA between the ears  $\leq$  15 dB. While all

participants were considered hearing aid candidates, none had worn hearing aids in the previous year. The participants were all native speakers of American English, had good self-reported health, and normal or corrected-to-normal vision ( $\leq 20/50$  on the Snellen Eye Chart), and passed the Montreal Cognitive Assessment with a score of 22 or better [14]

All study procedures were approved by the Institutional Review Boards of Northwestern University and the University of Colorado-Boulder. Participants completed an informed consent process and were compensated for their participation.

### Clinical Trial Design and Hearing Aid Parameter Settings

Full details regarding the clinical trial design can be found in Souza et al. [12]. Briefly, this trial was a double-blinded randomized crossover design. Participants attended a total of 8 visits to the laboratory. Each participant was fit with the same commercial hearing aids twice, once with a mild signal processing (MildSP) fitting and once with a strong signal processing (StrongSP) fitting. The order of fitting was randomized across participants, leading to some participants with an order of fitting MildSP -> StrongSP and others with an order of StrongSP -> MildSP. A licensed clinical audiologist completed the hearing aid fittings, follow-up visits, and follow-up phone calls. An experimenter collected all of the baseline and outcome measures. Participants wore each hearing aid condition for an average of 5-6 weeks. Participants were required to wear the hearing aids for at least 5 hours each day. Datalogging collected from the hearing aid fitting software was used to assess compliance with required wear time.

The hearing aids used in this study were premium-level commercial behind-the-ear (BTE) devices fit using a slim tube and custom earmold and were the most current

technology as of October 2015. Gain was fit using real ear verification and the NAL-NL2 prescriptive fitting rationale [15] and was matched to within 3 dB of targets under 2000 Hz and within 5 dB 2000-6000 Hz. Regardless of the fitting, gain was always set using the MildSP fitting, and if the participant was to receive the StrongSP condition first, the programming was then changed.

Using experimental software for both hearing aid fittings, the device parameter settings included omnidirectional microphones and disabled noise reduction. In the MildSP condition WDRC time constants were set to slow (attack: 1160 ms, release: 6900 ms) and nonlinear frequency compression (NLFC) was disabled. In the StrongSP condition the WDRC time constants were set to fast (attack: 13 ms, release: 59 ms) and NLFC was enabled. Every participant was given the same amount of NLFC (cutoff frequency set to 1.9 kHz and CR set to 3:1). At the follow-up visit, for participants who could not tolerate the initial gain settings, gain was decreased to the extent possible while staying with the tolerated range of deviation from targets. For those that were not able to tolerate the programmed NLFC settings, there was an allowed modification of increasing the cutoff frequency to 2.1 kHz and reducing the CR to 2:1.

### Self-report Measures and Collection

To study the subjective outcomes of hearing aid signal processing, we analyzed responses on the SSQ and the EAR. The SSQ is divided into three subscales. The first subscale has 14 items addressing speech perception. Questions on this subscale include items regarding speech understanding in quiet and in noise, and speech stream processing and switching. The second subscale, has 17 items addressing spatial hearing, including items querying the listener's ability to localize and perceive distance

and movement. The final subscale has 18 items addressing speech quality. These questions ask the listener to report on speech quality and naturalness, identify sounds and objects, segregate sounds, and listening effort. The SSQ uses a scale from 0 (least) -10 (greatest). Mean scores for each subscale and an overall score reflecting the average score across the three subscales are reported. The EAR scale is divided into two main subscales, each with two components: the Inner EAR subscale (pre- and post-fitting; hearing-related function) and the Outer EAR subscale (post-fitting only; hearing-aid-related function). The EAR subscales each have two components. The first is a global question which assesses overall perceived performance. The second is the survey (9 questions), which queries perceived performance on specific aspects of hearing-related function and device usage. The Inner EAR asks a listener to rate their ability to understand speech in quiet and noise as well as their emotional response to hearing. The Outer EAR includes questions regarding aided sound quality and naturalness, aided sound averseness, and device management. The EAR uses a 100 point scale with 0 the least and 100 the greatest for all questions. Mean scores for each EAR component are reported.

The SSQ and Inner EAR were collected at three time points: baseline, outcomes for session 1 and outcomes for session 2. The Outer EAR was completed at the two outcome measure visits. The SSQ self-report measure was completed on paper with the participants reading and answering questions on their own and the results transcribed by the experimenter to a spreadsheet, where scores were calculated. The Inner EAR and Outer EAR were completed online. That is, the EAR self-report measures were transcribed into a survey item within the electronic study database. Participants

completed the self-report measures in the laboratory by reading and answering questions on their own and results were stored and scores calculated within the database.

## **Results**

Statistical analyses were carried out using linear mixed effects (LME) models, implemented in R using the *lmer()* function from the *lme4* package [17-18]. The dependent variable was an average rating obtained on the SSQ or EAR self-report measure. Each SSQ/EAR subscale was analyzed in a separate model. The fixed factors included fitting (MildSP vs. StrongSP), session (Session 1 vs. Session 2), and order of fitting (MildSP → StrongSP vs. StrongSP → MildSP). Regression coefficients (*b*) were computed for each of these fixed factors with MildSP, Session 1, and MildSP → StrongSP order as the respective reference conditions. If baseline scores were available for a certain subscale, they were entered as covariates in the model. Participants nested within order were included as a random intercept. Regression coefficients indicated the direction of the effect and provided an estimate of the effect size.

## **Signal Modification**

[figure 2 and figure 3 near here]

Tables 1 and 2 summarize the LME analyses for the SSQ and EAR self-report measures, respectively. There was a significant effect of fitting on most of the subscales for both self-report measures. Among the SSQ subscales, there was a significant effect of fitting on the speech and quality scores, but not on the spatial scores. There was also a significant effect of fitting on the overall SSQ scores. Similarly, for the EAR self-report measure, there was a significant effect of fitting on the Inner EAR Global, Inner EAR

Survey, and Outer EAR Global scores, but not on the Outer EAR Survey scores. Figures 2 and 3 show that participants assigned a higher subjective rating to the MildSP condition than the StrongSP condition on the SSQ and EAR subscales, respectively.

[table 1 and table 2 near here]

Regression coefficients in Table 1 show that the overall effect of fitting on the SSQ outcomes was small ( $b < 1$ ). For instance, the change in signal modification from MildSP to StrongSP resulted in a decrease in the quality subscale score of 0.76 points. The effect on the speech subscale score was smaller. Regression coefficients in Table 2 show that the relative effect of fitting was larger on the EAR self-report measure. A change in fitting from MildSP to StrongSP resulted in a decrease in the Outer EAR Global subscale score by 15.76. Among the EAR subscales, the smallest decrease in scores was obtained for the Inner EAR Survey subscale by 9.45 points. The difference in beta values between the two self-report measures was most likely due to differences in the scales, i.e., SSQ scores range from 0-10, whereas EAR scores range from 0-100.

#### Order, Session, and Baseline

Effects of order and session were smaller than the effect of fitting on both self-report measures. There was a significant effect of order on the SSQ quality subscale and the Inner EAR Survey subscale. In both cases, the regression coefficients for this effect were positive, indicating that participants who received the StrongSP -> MildSP order of fitting assigned a higher rating than participants who received the MildSP -> StrongSP order of fitting. There was a significant effect of session only on the SSQ spatial subscale, with a lower rating for the second session compared to the first session.



Finally, there was a significant effect of the baseline score on all of the SSQ subscales and the Inner EAR Survey subscale, suggesting that higher baseline scores were associated with higher ratings after exposure to hearing aid signal processing. Figures 2 and 3 provide a visual comparison of the baseline scores to the fitting conditions for each subscale. It can be seen that the average Inner EAR Global and Inner EAR Survey scores at both levels of fitting were greater than the average baseline scores, suggesting an improvement in the aided condition. A similar trend can be seen for the SSQ speech subscale. However, the opposite trend was seen for the SSQ quality subscale, where in the average scores at both levels of fitting were poorer than the average baseline scores, suggesting a decrement in the aided condition.

## **Discussion**

The documentation of the subjective listener experience with amplification is vital in both the clinical and laboratory settings. The use of self-report measures may provide a window into the real-world experiences of an individual wearing hearing aids. By measuring perceived performance prior to a hearing aid fitting and then again following any changes to device signal processing parameters, it is possible to assess in a systematic way how listener impression changes as a function of device manipulation.

The results of the analyses presented here indicate that there was an overall positive effect of amplification, as shown by the improvements over baseline (unaided) on the Inner EAR module and the SSQ subscales. Additionally, the self-report measures considered here were able to capture the changes in perception when signal processing parameters were modified. In addition, the present results affirm that listener perception can and does change as a function of signal processing parameter manipulations.

The two measures we used had different sensitivities to changes in signal processing parameters. While significant, the sensitivity of SSQ seems to be lower than the sensitivity of EAR. If we consider the minimum clinically important difference (MCID) for EAR to be 6 points [19], then we see values higher than that between the two fits. Additionally, Yueh et al. [5] showed that the effect size for Inner EAR and Outer EAR scales is greater than other scales such as HHIE and APHAB. On the other hand, SSQ beta values were  $< 1$  in our study. Other studies have reported that at least a one point difference is needed on the SSQ in order to be considered a significant change [4,20]. For documenting changes in perception related to signal processing parameter settings, the EAR may provide a more sensitive measure for evaluating the impact of signal processing parameter modifications on listeners' self-reported perception.

Additionally, we note that the two measures used here may be sensitive to similar aspects of the hearing process, including speech perception in complex environments. For example, the Inner EAR components showed the same trend as the SSQ speech subscale. Both of these measures query aspects of speech intelligibility and we would expect that there would be similar responses within a listener. In contrast, the Outer EAR survey scale did not show the same degree of sensitivity to changes in signal processing modifications. This difference across survey scales may be because some of the questions on the scale had to do with hearing aid physical comfort, convenience, and dependability. These were not questions that would be affected by changes in signal processing, and would be expected to be the same across both the StrongSP and the MildSP fitting conditions. Responses to those questions may have been averaged out in the final cumulative ratings provided by the listeners.

The findings from this study reveal that the SSQ Spatial subscale did not show a difference for the StrongSP and the MildSp fittings. While both WDRC and NLFC may negatively impact the acoustic cues used in sound localization [e.g.,21-22], the amount of signal processing modifications selected for our two fitting conditions may not be sufficiently different from each other to differentially impact spatial sound perception. Another possible explanation is that spatial differences were not noticeable enough in every day listening environments so as to be captured by the post-fitting self-report measures. The use of alternative subjective assessments (e.g., ecological momentary assessment (EMA)) [23] may identify facets of perception within a listener's environment that are not sufficiently different to be recalled when subjective ratings are provided after the fact.

The current study has several limitations. The hearing aid signal processing parameters used in the NLFC in the StrongSP condition are beyond typical clinical implementation of NLFC given the degree and configuration of hearing loss of these participants. This was a useful difference in this case, since we wished to be sure acoustically different conditions were being compared. However, we would expect to see larger differences in our participants' ratings than a clinician may see if more subtle processing differences had been implemented. Our goal in selecting processing parameters was to assure that the parameter modifications made substantial changes to the output of the hearing aid, and in that aim, we were successful. In addition, the participants in this study were new hearing-aid users. While these factors may limit the generalizability of our findings to all clinical populations, the results support the use of self-report measures as effective tools in documenting the perceptual responses of

listeners to hearing aid fittings that vary substantially in the settings of the signal processing features.

## **Conclusion**

This study examined whether self-perceived performance with hearing aids was sensitive to differences in signal processing. Using the SSQ and the EAR, we showed that participants self-rate their perceptions differently for a hearing aid fitting that implemented only a mild amount of signal processing compared to a fitting with the same device that implemented stronger processing settings. Consistent with the idea that published surveys query patients on different aspects of their listening experiences, the two instruments considered here (SSQ and EAR) were sensitive to different facets of perceived benefit.

## **Acknowledgements**

This project was supported by NIH grant R01 DC012289 to P. Souza and K. Arehart. We thank Cynthia Erdos, Laura Mathews, Elizabeth McNichols, Arianna Mihalakakos, Kristin Sommerfeldt, and Melissa Sherman for their assistance with data collection and management; and Christine Jones and Olaf Strelcyk for project support. The project was a registered NIH clinical trial (ClinicalTrials.gov Identifier: NCT02448706). Data management via REDCap is supported at Feinberg School of Medicine by the Northwestern University Clinical and Translational Science (NUCATS) Institute. Research reported in this publication was supported, in part, by the National Institutes of Health's National Center for Advancing Translational Sciences, Grant Number

UL1TR001422. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

**Disclosure**

No conflict of interest is reported by the authors.

## References

1. American Academy of Audiology Task Force (2006). Guideline for audiologic management of the adult patient. *Audiology Online*.
2. Cox, R. M., & Alexander, G. C. (1995). The abbreviated profile of hearing aid benefit. *Ear and Hearing, 16*(2), 176-186.
3. Cox, R. M., & Alexander, G. C. (2000). Expectations about hearing aids and their relationship to fitting outcome. *Journal of the American Academy of Audiology, 11*(7).
4. Noble, W., & Gatehouse, S. (2006). Effects of bilateral versus unilateral hearing aid fitting on abilities measured by the Speech, Spatial, and Qualities of Hearing scale (SSQ). *International Journal of Audiology, 45*(3), 172-181.
5. Yueh, B., McDowell, J. A., Collins, M., Souza, P. E., Loovis, C. F., & Deyo, R. A. (2005). Development and validation of the effectiveness of auditory rehabilitation scale. *Archives of Otolaryngology–Head & Neck Surgery, 131*(10), 851-856.
6. Abrams, H. B., Chisolm, T. H., McManus, M., & McArdle, R. (2012). Initial-fit approach versus verified prescription: comparing self-perceived hearing aid benefit. *Journal of the American Academy of Audiology, 23*(10), 768–778.
7. Shi, L.-F., Doherty, K. A., Kordas, T. M., & Pellegrino, J. T. (2007). Short-term and long-term hearing aid benefit and user satisfaction: a comparison between two fitting protocols. *Journal of the American Academy of Audiology, 18*(6), 482–495.
8. Dawes, P., Munro, K. J., Kalluri, S., & Edwards, B. (2014). Acclimatization to hearing aids. *Ear and Hearing, 35*(2), 203–212.
9. Souza P, Yueh B, McDowell J, Collins M, Loovis C, Boike K, Deyo R. (2002). Sensitivity of self assessment questionnaires to differences in hearing aid technology. International Hearing Aid Conference, Tahoe City, CA.
10. Alcántara, J. I., Moore, B. C. J., & Marriage, J. (2004). Comparison of three procedures for initial fitting of compression hearing aids. II. Experienced users, fitted unilaterally. *International Journal of Audiology, 43*(1), 3–14.
11. Anderson, M.C., Arehart, K.H., & Souza, P.E. (2018). Survey of Current Practice in the Fitting and Fine-Tuning of Common Signal-Processing Features in Hearing Aids for Adults, *Journal of the American Academy of Audiology, 29*(2), 118-124.
12. Souza, P., Arehart, K., Schoof, T., Anderson, M., Stori, D., and Balmert, L. (submitted). Understanding variability in individual response to hearing aid signal processing: Results from a clinical trial. *Submitted for publication*.
13. Gatehouse, S., & Noble, W. (2004). The speech, spatial and qualities of hearing scale (SSQ). *International Journal of Audiology, 43*(2), 85-99.

14. Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J.L. & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695-699.
15. Keidser, G., Dillon, H., Flax, M., Ching, T., & Brewer, S. (2011). The NAL-NL2 prescription procedure. *Audiology Research*, 1(1).
16. Gatehouse, S., & Akeroyd, M. (2006). Two-eared listening in dynamic situations: Audición con dos oídos en situaciones dinámicas. *International Journal of Audiology*, 45(sup1), 120-124.
17. Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. doi:10.18637/jss.v067.i01.
18. R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
19. Collins, M. P., Liu, C. F., Taylor, L., Souza, P. E., & Yueh, B. (2013). Hearing aid effectiveness after aural rehabilitation: Individual versus group trial results. *Journal of rehabilitation research and development*, 50(4), 585-598.
20. Lenarz, T., Muller, L., Czerniejewska-Wolska, H., Varela, H. V., Dotú, C. O., Durko, M., ... & Graham, P. L. (2017). Patient-Related Benefits for Adults with Cochlear Implantation: A Multicultural Longitudinal Observational Study. *Audiology and Neurotology*, 22(2), 61-73.
21. Brown, A. D., Rodriguez, F. A., Portnuff, C. D., Goupell, M. J., & Tollin, D. J. (2016). Time-varying distortions of binaural information by bilateral hearing aids: effects of nonlinear frequency compression. *Trends in hearing*, 20, 2331216516668303.
22. Hassager, H. G., Wiinberg, A., & Dau, T. (2017). Effects of hearing-aid dynamic range compression on spatial perception in a reverberant environment. *The Journal of the Acoustical Society of America*, 141(4), 2556-2568.
23. Wu, Y. H., Stangl, E., Zhang, X., & Bentler, R. A. (2015). Construct validity of the ecological momentary assessment in audiology research. *Journal of the American Academy of Audiology*, 26(10), 872-884.

**Table 1.** Results for linear mixed effects models for the effects of fitting, session, and order on average ratings obtained for each subsection and overall scores of the Speech, Spatial, and Quality Questionnaire (SSQ). Baseline scores were controlled for in each model. Regression coefficients (*b*), standard error (*SE*), and 95% confidence intervals for the regression coefficients (*b* 95% CI; LL = lower limit, UL = upper limit) are provided for each factor. Significant effects are represented in bold with asterisks (\**p*<0.05, \*\**p*<0.01, \*\*\**p*<0.001).

		<b>SSQ</b>			
		<b>Speech</b>	<b>Spatial</b>	<b>Quality</b>	<b>Overall</b>
<b>Modification</b>	<i>b</i>	<b>-0.641**</b>	-0.370	<b>-0.756***</b>	-0.580
	SE	0.216	0.191	0.176	0.148
	<i>p</i>	0.005	0.061	< 0.001	< 0.001
	<i>b</i> 95% CI [LL,UL]	[-1.064, -0.217]	[-0.744, 0.004]	[-1.100, -0.411]	[-0.870, -0.289]
<b>Session</b>	<i>b</i>	-0.019	<b>-0.421*</b>	-0.043	-0.164
	SE	0.214	0.1884	0.174	0.146
	<i>p</i>	0.929	0.029	0.804	0.264
	<i>b</i> 95% CI [LL,UL]	[-0.438, 0.400]	[-0.791, -0.052]	[-0.384, 0.297]	[-0.451, 0.123]
<b>Order</b>	<i>b</i>	0.579	0.548	<b>0.806*</b>	0.590
	SE	0.374	0.423	0.345	0.323
	<i>p</i>	0.129	0.201	0.046	0.085
	<i>b</i> 95% CI [LL,UL]	[-0.154, 1.312]	[-0.281, 1.378]	[0.130, 1.483]	[-0.043, 1.223]
<b>Baseline</b>	<i>b</i>	<b>0.378**</b>	<b>0.292*</b>	<b>0.501***</b>	0.483
	SE	0.113	0.120	0.106	0.120
	<i>p</i>	0.002	0.019	< 0.001	1
	<i>b</i> 95% CI [LL,UL]	[0.157, 0.598]	[0.058, 0.526]	[0.294, 0.708]	[0.248, 0.718]



**Table 2.** Results for linear mixed effects models for the effects of fitting, session, and order on the average ratings from the Effectiveness of Auditory Rehabilitation (EAR) scale. Baseline scores were controlled only for the Inner EAR subsections of the scale. Regression coefficients (*b*), standard error (*SE*), and 95% confidence intervals for the regression coefficients (*b* 95% CI; LL = lower limit, UL = upper limit) are provided for each factor. Significant effects are represented in bold with asterisks (\**p*<0.05, \*\**p*<0.01, \*\*\**p*<0.001).

		<b>EAR</b>			
		<b>Inner Global</b>	<b>Inner Survey</b>	<b>Outer Global</b>	<b>Outer Survey</b>
<b>Modification</b>	<i>b</i>	<b>-14.044**</b>	<b>-9.495***</b>	<b>-15.768***</b>	-2.208
	SE	4.297	2.591	4.152	2.372
	<i>p</i>	0.003	<0.001	< 0.001	0.362
	<i>b</i> 95% CI [LL,UL]	[-22.465, -5.623]	[-14.572, -4.418]	[-23.905, -7.631]	[-6.856, 2.441]
<b>Session</b>	<i>b</i>	-0.480	-0.820	2.045	0.316
	SE	4.269	2.570	4.119	2.347
	<i>p</i>	0.911	0.750	0.620	0.893
	<i>b</i> 95% CI [LL,UL]	[-8.848, 7.887]	[-5.857, 4.216]	[-6.028, 10.119]	[-4.284, 4.915]
<b>Order</b>	<i>b</i>	4.414	<b>8.288*</b>	10.878	0.108
	SE	5.016	3.571	5.662	4.060
	<i>p</i>	0.381	0.047	0.074	0.979
	<i>b</i> 95% CI [LL,UL]	[-5.417, 14.245]	[1.290, 15.287]	[-0.219, 21.974]	[-7.850, 8.066]
<b>Baseline</b>	<i>b</i>	0.219	<b>0.524***</b>	N/A	N/A
	SE	0.170	0.113		
	<i>p</i>	0.203	< 0.001		
	<i>b</i> 95% CI [LL,UL]	[-0.114, 0.552]	[0.301, 0.746]		

## Figure Captions

Figure 1. Audiograms of participants. Thick black line represent average audiogram for each ear.

Figure 2. Box-and-whisker plots for the effect of fitting (MildSP vs. StrongSP) on x-axis for average ratings on the Speech, Spatial, and Qualities (SSQ) Questionnaire (y-axis) across # participants. Panels a-c represent speech, spatial, and qualities subscales, and panel d represents the overall average ratings across the three subscales. The black dashed line represents the average baseline rating.

Figure 3. Box-and-whisker plots for the effect of fitting (MildSP vs. StrongSP) on x-axis for average ratings on the Effectiveness of Auditory Rehabilitation (EAR) scale (y-axis) across # participants. Panels a-c represent Inner EAR (Global), Inner EAR (Survey), Outer EAR (Global), and Outer EAR (Survey) subscales. The black dashed line represents the average baseline rating for the Inner EAR subscales.

Figure 1.

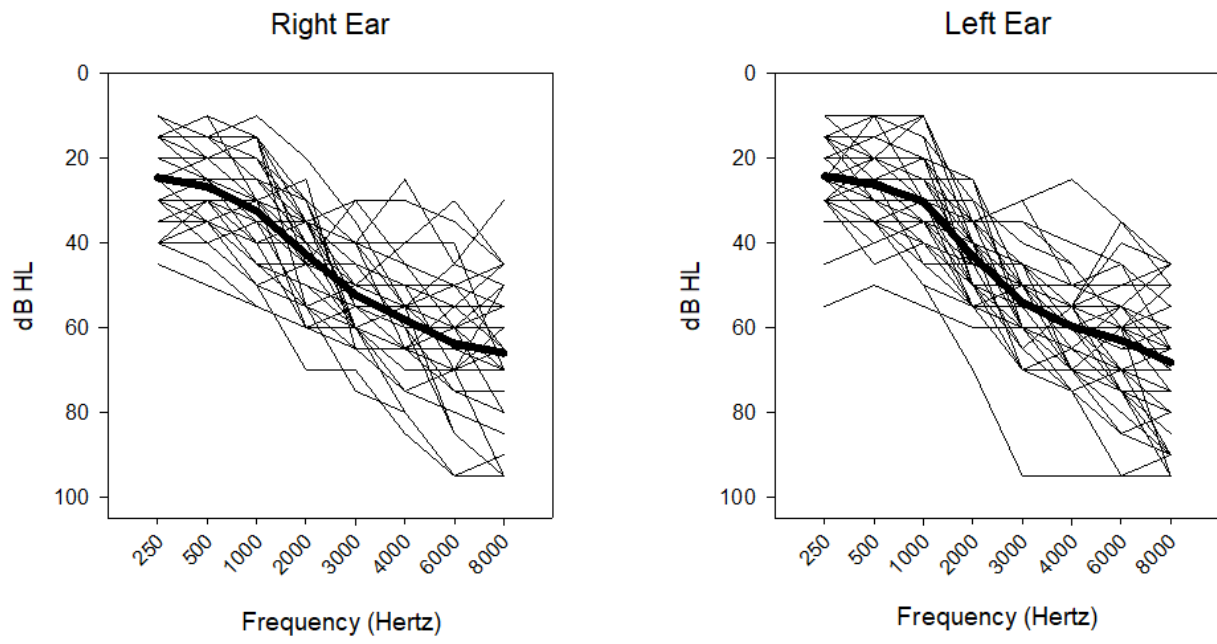


Figure 2.

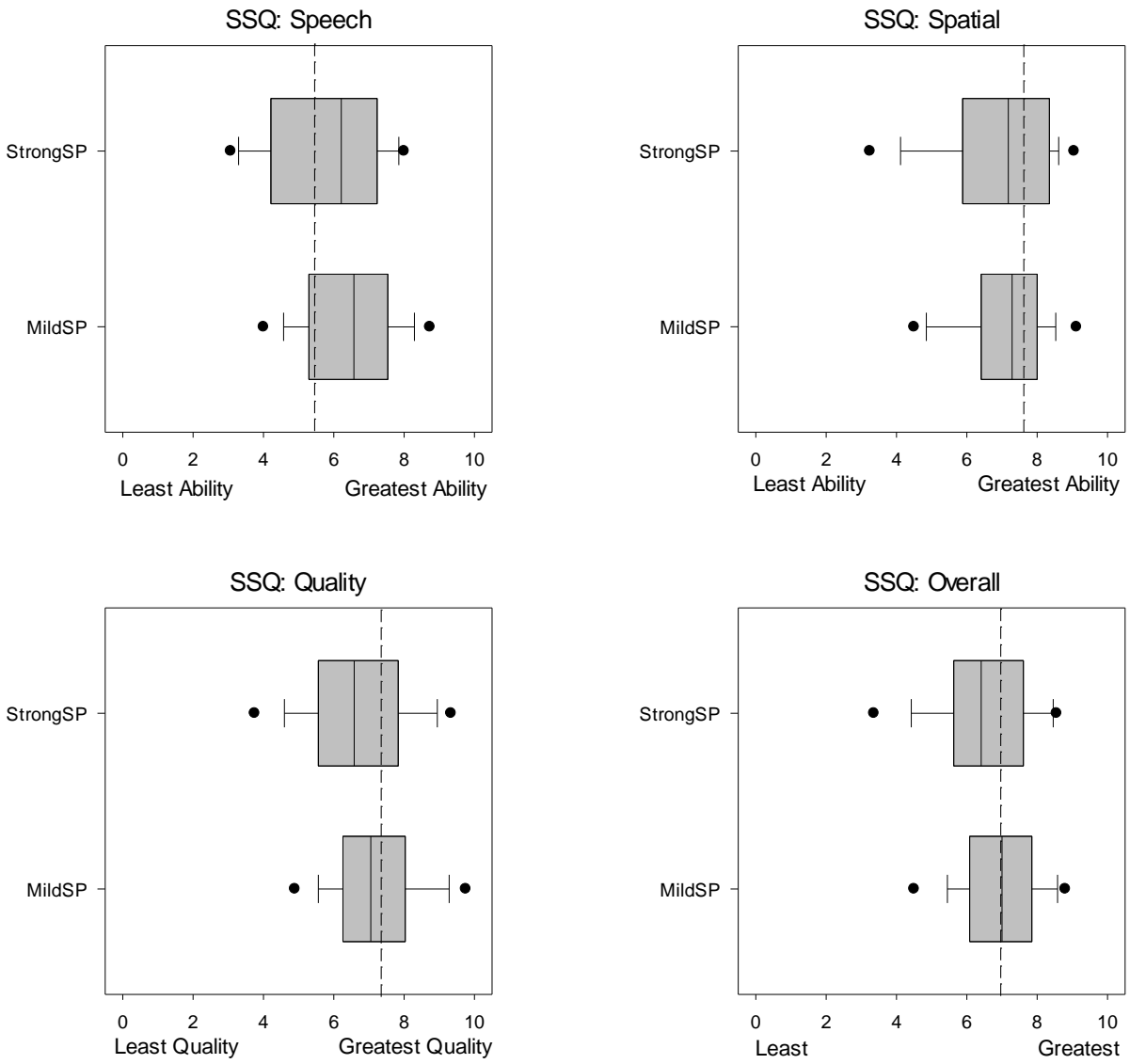


Figure 3.

