

# Hearing Aid Self-Adjustment: Effects of Formal Speech-Perception Test and Noise

Trends in Hearing  
Volume 24: 1–16  
© The Author(s) 2020  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/2331216520930545  
journals.sagepub.com/home/tia



Carol L. Mackersie<sup>1</sup> , Arthur Boothroyd<sup>1,2</sup> , and Harinath Garudadri<sup>2</sup>

## Abstract

While listening to recorded sentences with a sound-field level of 65 dB SPL, 24 adults with hearing-aid experience used the “Goldilocks” explore-and-select procedure to adjust level and spectrum of amplified speech to preference. All participants started adjustment from the same generic response. Amplification was provided by a custom-built Master Hearing Aid with online processing of microphone input. Primary goals were to assess the effects of including a formal speech-perception test between repeated self-adjustments and of adding multitalker babble (signal-to-noise ratio +6 dB) during self-adjustment. The speech test did not affect group-mean self-adjusted output, which was close to the National Acoustics Laboratories’ prescription for Non-Linear hearing aids. Individuals, however, showed a wide range of deviations from this prescription. Extreme deviations at the first self-adjustment fell by a small but significant amount at the second. The multitalker babble had negligible effect on group-mean self-selected output but did have predictable effects on word recognition in sentences and on participants’ opinion regarding the most important subjective criterion guiding self-adjustment. Phoneme recognition in monosyllabic words was better with the generic starting response than without amplification and improved further after self-adjustment. The findings continue to support the efficacy of hearing aid self-fitting, at least for level and spectrum. They do not support the need for inclusion of a formal speech-perception test, but they do support the value of completing more than one self-adjustment. Group-mean data did not indicate a need for threshold-based prescription as a starting point for self-adjustment.

## Keywords

hearing loss, hearing aids, self-fitting, self-adjustment, amplification

Received 20 July 2019; Revised 5 April 2020; accepted 1 May 2020

Untreated hearing loss is associated with a number of comorbidities including an increased risk of falls, dementia, social isolation, and depression (Abrams, 2017; Besser et al., 2018). Despite these risks and the impact of untreated hearing loss on quality of life (Brodie & Ray, 2018; Hogan et al., 2009), a substantial percentage of hearing loss remains untreated (Lin et al., 2011). There are many factors underlying low hearing-aid uptake. One of the most frequently cited reasons is the perception that the hearing problem is not bad enough to warrant treatment (Knudsen et al., 2010; Powers & Rogin, 2019; Tahden et al., 2018). In a recent survey conducted in the United States, the high cost of hearing aids was the second most frequently cited reason for not acquiring hearing aids (Powers & Rogin, 2019). The recent passage of the Over-the-Counter Hearing Aid Act (2017) is expected to ease the obstacles to acquiring hearing aids by enabling persons with

self-identified mild-moderate hearing loss to obtain hearing amplification “without the supervision, prescription, or other order, involvement, or intervention of a licensed person” (p. 1). The successful use of over-the-counter (OTC) aids, however, will require the consumer to independently assemble, adjust, and operate the aids without professional assistance. This study is one of a series addressing self-adjustment of hearing aids as one component in the self-fitting process.

<sup>1</sup>School of Speech, Language and Hearing Sciences, San Diego State University

<sup>2</sup>Qualcomm Institute of Calit2, University of California, San Diego

### Corresponding Author:

Carol L. Mackersie, School of Speech, Language and Hearing Sciences, San Diego State University, San Diego, CA, United States.

Email: cmackers@sdsu.edu



At the time of writing, the implementation of the OTC hearing aids awaits the Food and Drug Administration guidelines that are needed to ensure quality and safety. There are, however, OTC amplification devices called “personal sound amplification products” (PSAPs) currently available. Although not approved to compensate for hearing loss, PSAPs have nevertheless been adopted by some people as a substitute for professionally fit hearing aids (Kochkin, 2010).

User self-adjustment options for currently available PSAPs vary from simple volume controls to controls that enable the selection of frequency response and variations of dynamic range compression (Almufarrij et al., 2019; Brody et al., 2018; Nelson et al., 2018). For people with mild-to-moderate hearing loss, speech recognition obtained with PSAPs has been shown to be better than with no amplification (Brody et al., 2018; Reed et al., 2017; Sacco et al., 2016) although exceptions have been reported for low-quality devices (Reed et al., 2017). In a recent study, for example, Brody et al. (2018) compared sentence recognition obtained with three different PSAPs to that obtained with a commercially-available hearing aid. One PSAP enabled user adjustment of the frequency response in addition to overall gain, whereas the other two only enabled overall gain adjustments with a volume control. Mean scores on the Hearing in Noise Test in quiet, Speech Intelligibility Index (SII) values (American National Standards Institute [ANSI], R2012), and listening effort ratings of the one PSAP that enabled user frequency-response adjustments were similar to the those obtained with the hearing aid. In contrast, the PSAPs with only a volume control had significantly lower SIIs than both the hearing aid and the PSAP with the option for user frequency-response adjustment. These results suggest that direct-to-consumer aids that include user adjustment of the frequency response may lead to better outcomes.

The efficacy of user adjustment of amplification has been demonstrated using a variety of protocols that incorporate changes of frequency response (Boothroyd & Mackersie, 2017; Dreschler et al., 2008; Jensen et al., 2019; Keidser & Convery, 2018; Keidser et al., 2008; Mackersie et al., 2019; Nelson et al., 2018). Group-mean user-adjusted frequency responses are generally close to threshold-based prescribed targets (typically within 5 dB), but substantial individual differences have been observed (Jensen et al., 2019; Keidser & Convery, 2018; Mackersie et al., 2019; Nelson et al., 2018; Punch et al., 1994). In addition, Nelson et al. (2018) reported that on average, speech recognition outcomes of self-adjusted responses were not significantly different from those using the current version of the National Acoustics Laboratories prescription for Non-Linear hearing aids (NAL-NL2; Keidser et al., 2011). No evidence has been found that individual differences

of self-adjustment relative to prescriptive targets are related to listener characteristics such as age, gender, average hearing loss, hearing-aid experience, or noise tolerance (Perry et al., 2019).

An alternative to self-fitting, based on direct manipulation of frequency response and other parameters, is to provide a limited set of fixed responses from which to choose. Using this approach, Humes et al. (2017) compared self-fitting outcomes to those obtained from a conventional audiologist’s fitting. Participants in the self-fitting (consumer decides) group tried up to three hearing aids programmed with three different frequency responses and chose the device they preferred. Mean self-perceived benefit as measured by the Profile of Hearing Aid Benefit (PHAB) was not significantly different for the self-fit and audiologist-fit groups, even though several participants did not choose the frequency response that was closest to prescribed, threshold-based, targets.

We previously reported on a self-fitting study using an explore-and-select procedure we named “Goldilocks” (Boothroyd & Mackersie, 2017; Mackersie et al., 2019). This study was implemented with preprocessed stimuli presented from a computer. In other words, there was no active microphone to provide auditory input during instruction or self-hearing. Participants self-adjusted overall output, high-frequency boost, and low-frequency cut to preference while listening to recorded sentences spoken by a man. The intent was to evaluate self-adjustments that would not rely on information about the user’s audiogram. Therefore, instead of starting from an individual prescribed response, as others have done, we started every participant with the same amplification designed to match the NAL-NL2 targets for a single, generic, mild-to-moderate hearing loss. The choice of audiogram from which to create a starting response was based on a report of the most common audiometric configurations (Ciletti & Flamme, 2008). Specifically, the thresholds selected to represent a mild-moderate hearing loss for this study closely matched the mean of the male and female thresholds for symmetrical hearing for the mildest configuration that had 2000 Hz thresholds of at least 30 dB HL.

The full Goldilocks protocol, used in the earlier study, included a formal speech-perception test after an initial self-adjustment and was followed by a second adjustment. The goal of this formal test was to increase listeners’ reliance on intelligibility as a self-adjustment criterion. Overall, 77% of participants were able to complete the self-adjustment protocol without assistance. After completing the self-adjustments, 88% reached an SII criterion of 0.6, which, on average, provided a 95% word-in-sentence recognition score. There was a significant increase of self-selected high-frequency output following (but not necessarily because of) the

speech-perception test. This effect, however, was restricted to participants with previous hearing-aid experience. Nonusers showed no significant change after taking the speech-perception test. A primary goal of this study was to determine the need for inclusion of a formal speech-perception test in the explore-and-select procedure. Additional goals were to assess individual and group outcomes of self-adjustment in quiet and in noise.

Whereas our previous study used preprocessed stimuli presented through an earphone, this study used a desktop Master Hearing Aid (MHA) developed for this project under a subcontract to the University of California, San Diego (Garudadri et al., 2017). The MHA used real-time wide-dynamic-range compression processing of input from an ear-level microphone. The output was delivered to a receiver in the ear canal.

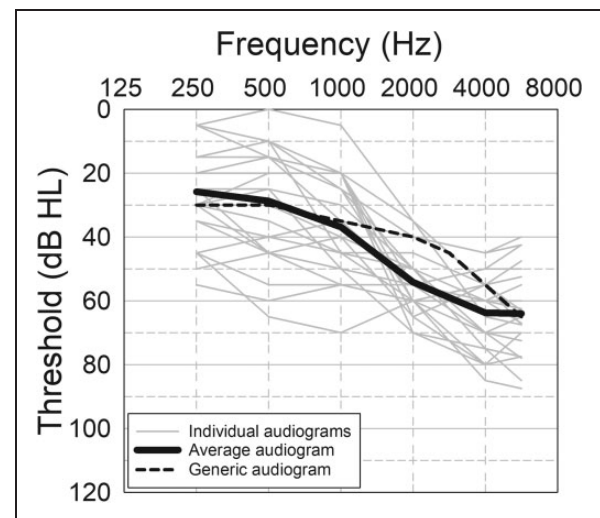
Specific goals were as follows:

1. To assess outcomes of self-fitting, using real-time processing of sound-field speech input to an active microphone, in terms of real-ear output level and spectrum, speech-perception performance, number of adjustment steps, and time taken.
2. To measure self-adjustment replication effects on real-ear output with and without an intervening speech-perception test.
3. To determine the effect of multitalker babble on the outcome of self-fitting.
4. To determine how group-mean and individual user-selected frequency responses compare with responses prescribed by NAL-NL2.
5. To determine the subjective criteria felt to be important to participants during self-adjustment.

## Methods

### Participants

Twenty-four adult hearing-aid users (12 men and 12 women) participated. This study was limited to aid-users because the previous study only found evidence of change between a first and second self-adjustment in hearing-aid users. These participants did not have any prior experience with self-adjustment studies. The sample size was based on power analyses conducted on data from the previous study that indicated a minimum sample of 18 participants was needed to detect group differences with a power goal of 0.80 and an  $\alpha$  level of .05. Age range was from 49 to 86 years with a mean of 72 years. All participants had a score on the Montreal Cognitive Assessment of 21 or higher (Nasreddine et al., 2005). Mean and individual audiograms of the better ear (used for testing) are shown in Figure 1 along with the generic audiogram used to define the starting response from which all self-adjustment would occur.



**Figure 1.** Audiograms of the 24 Participants' Test Ears. The heavy line shows the mean audiogram. The dashed line shows the audiogram used to create the starting response for all participants.

**Table 1.** Mean Age, 4FA Hearing Loss, Years of HA exp, MoCA Scores, and Ed yrs.

Variable	Control		Experimental		<i>t</i>	<i>p</i>
	Mean	<i>SD</i>	Mean	<i>SD</i>		
Age	74.0	7.7	70.2	8.2	1.18	0.25
4FA	46.5	9.4	45.3	11.4	0.27	0.79
HA exp	16.8	28.4	14.5	28.0	0.20	0.84
MoCA	26.3	3.2	27.1	2.6	-0.69	0.50
Ed yrs	16.1	3.0	17.2	2.9	-0.90	0.38

Note. 4FA = four-frequency average; HA exp = years of hearing aid experience; MoCA = Montreal Cognitive Assessment; Ed yrs = years of education; *SD* = standard deviation.

The research was approved by the San Diego State University Institutional Review Board. All participants signed a written consent before any data were collected.

To address the value of formal speech-perception testing within the self-fitting procedure, participants were randomly assigned to two groups. An "exposure" group received speech recognition testing between a first and second self-adjustment. A control group performed a nonauditory task instead. Table 1 shows descriptive statistics for the two groups, with no evidence of significant differences. Both groups completed the two self-adjustments in quiet and in a background of spectrally matched multitalker babble using a signal-to-noise ratio (SNR) of +6 dB. This was a mixed design with exposure group (exposed vs. control) as a between-subjects variable and replications (Adjustment 1 vs. Adjustment 2) and noise (quiet vs. multitalker babble) as within-subjects variables. The exposure group

completed the nonauditory task after the second self-adjustment and the control group completed the speech-perception test. This was done to ensure both groups had the same total experience before speech-perception outcome assessment.

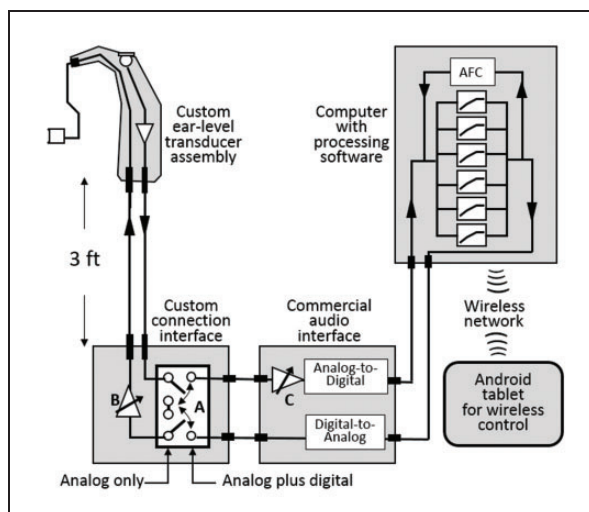
### Master Hearing Aid

**MHA Hardware.** The MHA developed for this study had five components:

1. A custom-built ear-level transducer assembly in a behind-the-ear case, receiver-in-canal, and built-in microphone pre-amplifier.
2. A custom-built analog interface including a power amplifier with adjustable gain.
3. A commercial audio interface, with adjustable input, for analog/digital and digital/analog conversion (Zoom Tac-8, Zoom, Hauppauge, New York).
4. A MacBook computer incorporating the custom-designed speech-processing software.
5. An Android tablet incorporating the Goldilocks self-fitting interface for wireless control of the processing software.

The five components are illustrated in Figure 2. The switches (A) and gain controls (B and C) allowed calibration of the analog and digital components so that decibel readings in the control software matched acoustic gain and 2cc-coupler-output values as closely as possible.

**MHA Software.** The processing software, adapted from a design by Kates (Souza et al., 2015), provided control of amplification in six bands. The finite impulse response

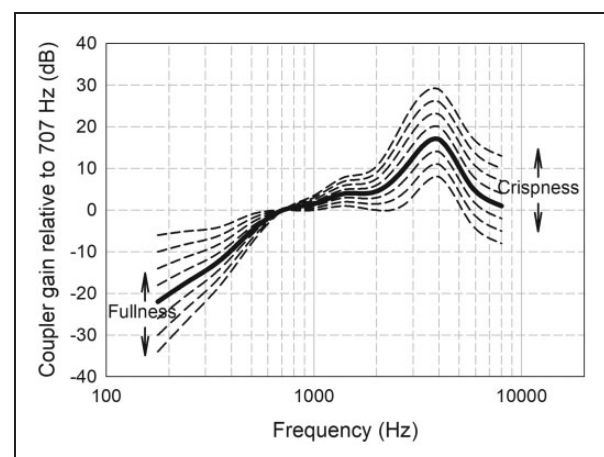


**Figure 2.** The Five Components of the University of California, San Diego Master Hearing Aid. AFC = automatic feedback control.

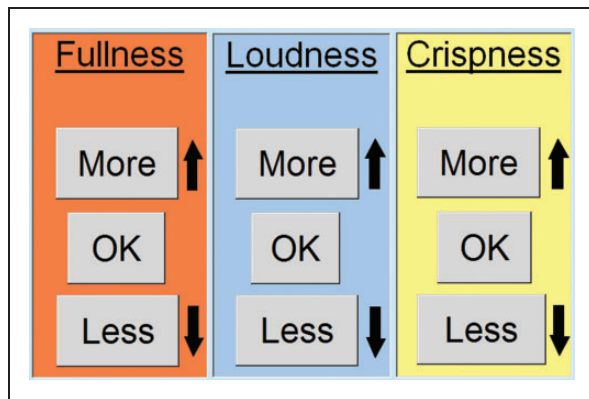
(FIR) filter cross-over frequencies used in this study were 250, 500, 1000, 2000, and 4000 Hz. The center frequencies were, approximately, 177, 354, 707, 1414, 2828, and 5657 Hz. As measured using ANSI 3.22 testing on an Audioscan Verifit 1 (ANSI, 2014), the equivalent input noise level was 29 dB SPL, and the frequency bandwidth was 200 to 8000 Hz. A researcher programming interface provided for adjustment of gains, compression ratios, compression thresholds, compression dynamics, and maximum outputs in each band. Automatic feedback control provided around 15 dB of added stable gain. The processing delay, between input and output, was measured at 8.2 ms. For this study, the compression ratio in each band was set to 1.4:1 with a threshold of 45 dB SPL. Nominal attack and release times were 20 and 100 ms. Maximum band output was set to 110 dB SPL. The Goldilocks programming interface also included selection of step sizes for overall gain, high-frequency boost, and low-frequency cut. For this study, overall gain was adjusted in 3 dB steps. The step sizes for high-frequency boost and low-frequency cut were adjusted differently in each frequency band but reached 3 dB at 250 Hz and 3 kHz, as shown in Figure 3.

### Goldilocks Software

The Goldilocks user interface, shown in Figure 4, was unchanged from that used in Mackersie et al. (2019). The underlying software, however, was adapted as an Android application for control of the processing platform via a WiFi link. The “Fullness” and “Crispness” controls provided for listener adjustment of low-frequency cut and high-frequency boost, as seen earlier in Figure 3 where the heavy line shows the generic starting response. The use of a frequency fulcrum around which gain changes were made is similar to the approach



**Figure 3.** Range of Responses Made Available to the Listener. The heavy line shows a generic starting response for self-adjustment.



**Figure 4.** The Three Controls Available to the User During Self-Adjustment. To help participants learn their effects, these controls were presented one-at-a-time at first, and then shown together for final adjustment.

used by Punch and Robb (1992). In that study, however, a single adjustment of slope was used, whereas, in this study, slopes above and below the fulcrum were adjusted independently.

### Speech Perception Measures

Two speech-perception measures were included in outcome assessment:

1. Phoneme recognition in words, using Boothroyd's Computer-Assisted Speech-Perception Assessment test (Boothroyd, 2008; Mackersie et al., 2001). This test eliminates effects of sentence context on word recognition but retains an effect of word context on phoneme recognition (Boothroyd, 1968, 2008; Boothroyd & Nittrouer, 1988).
2. Word recognition in sentences using the City University of New York (CUNY) sentence test (Boothroyd et al., 1988; Hanin et al., 1988). This test provides high levels of semantic and syntactic context.

Sentences from the CUNY test were also used as the material heard during self-adjustment, but different lists were used for self-adjustment and outcome assessment. In addition, the CUNY test was used as the formal speech-perception test between the first and second self-adjustments for the exposure group and after the second self-adjustment for the control group.

### Speech-Intelligibility Index

Individual estimates of speech-weighted audibility were obtained for a speech input of 65 dB (SPL root mean square [RMS]) using ANSI Revision 2007 and a frequency-importance function derived from the NU6

data in Table B.2 (ANSI, 1997). The result was an estimate of the proportion of the useful speech information, in a signal presented at a sound-field level of 65 dB SPL, that was audible to the listener. SII estimates in quiet were obtained for the generic starting response, the two self-adjustments and the NAL-NL2 targets.

### Adjustment-Criteria Questionnaire

After each of four self-adjustments (two in quiet and two in multitalker babble), participants completed an adjustment-criterion questionnaire designed for this study. Using a 5-point scale, participants rated the importance of four subjective criteria: (a) loudness (not too loud or quiet), (b) clarity (easy to understand), (c) quality (natural—similar to what I expect to hear), and (d) noisiness (bothered by noise). The end points were *not at all important* (1) and *very important* (5). For consistency, the Noisiness Scale was included, even when testing without babble. After completing the absolute 5-point ratings, participants were asked to select one of the factors as being the most important.

### Procedure

**Aid Adjustment.** Self-adjustment and testing were monaural using the ear with the better threshold at 2 kHz. The ear was chosen randomly if there was no difference. To minimize acoustic feedback problems, closed Oticon power domes were used to couple the hearing-aid receiver to the ear. The participants wore the ear-level assembly during both self-adjustment and testing. The nontest ear was occluded with a foam earplug for participants who had nontest ear thresholds of 40 dB HL or less at any frequency.

All adjustments were made in a double-walled sound booth while listening to concatenated CUNY sentences recorded by a woman and presented in the sound field at a level of 65 dB SPL (at the listener's location) from a loudspeaker positioned 3 feet in front of the listener. Although these materials were prerecorded, they were presented in the sound field to provide microphone input.

As indicated earlier, all participants started adjustment from a frequency response matched to the NAL-NL2 prescription for the generic mild-to-moderate sloping hearing loss shown in Figure 1. The frequency response was matched only for a 65 dB SPL input speech signal on the Verifit. Therefore, there was no verification that frequency responses at other input levels matched the NAL-NL2 compression parameters. The three parameter controls shown in Figure 3 were initially presented one at a time in the following order: (a) overall level, (b) high-frequency boost, (c) repeat overall level (optional), and (d) low-frequency cut. Participants were instructed to adjust each sound parameter by

increasing until the sensation was “too much” and by decreasing until it was “too little” before finding the value that was “just right.” These instructions were intended to ensure that participants would explore the full range of acceptability for each setting. Following self-adjustment of the controls in isolation, the three were presented together and participants were given the option of making final adjustments to all three parameters as shown in Figure 3.

All participants completed the self-adjustments separately in quiet and in multitalker babble (+6 dB SNR). The quiet and noise adjustments were performed on 2 separate days, the order of which was counterbalanced across participants.

After completing a first self-adjustment, half of the participants (the “exposure” group) took a word-in-sentence recognition test (CUNY) under the listening condition just used for self-adjustment (quiet or noise) and without feedback on performance. The recognition test provided goal-directed listening experience with the aid as they had just adjusted it. After a second self-adjustment (starting with the same generic response), they performed a 4-min nonauditory task involving the manipulation of shapes on a Samsung tablet. Participants continued wearing the MHA during the task, but there was no talking during the task and no external sounds were played. The nonauditory task took approximately the same duration as that required to complete the recognition test administered between the first and second adjustments. The other half of the participants (the “control” group) performed the nonauditory task between the two self-adjustments and the speech-perception task after the second. This arrangement allowed assessment of the effect of the sentence-recognition test on the second response adjustment by the exposure group while ensuring, as indicated earlier, that both groups had the same total experience before final speech-perception outcome measurement.

**Outcome Measures.** Outcome of self-fitting was assessed by five measures:

**Real-Ear Measures.** Using the Verifit 1 hearing-aid analyzer, real-ear output data were obtained for the individual generic starting response and for the four self-adjustments (two in quiet and two in multitalker babble) using a 65 dB SPL female speech input (the “carrot” story). NAL-NL2 targets were also obtained using individual participants’ thresholds.

**Speech Recognition.** After completion of two self-adjustments in quiet (plus the formal speech-perception test and the nonauditory task), phoneme-in-word recognition was measured in quiet, at four levels (45, 55, 65, and 75 dB SPL), for three conditions: unaided, generic

response, and final self-adjustment. In addition, word-in-sentence recognition was measured at a speech presentation level of 60 dB SPL, both in quiet and with a speech-to-babble ratio of +6 dB.

**Self-Adjustment Criteria.** Immediately after completing each self-adjustment (two in quiet and two in noise), participants were asked to rate each of the four adjustment criteria in terms of importance and, also, to select one as being the most important.

**Number of Adjustments Made, and Time Taken.** These data were logged automatically by the MHA software.

**Abbreviated Profile of Hearing aid Benefit—Aversiveness Subscale.** In addition to the outcome measures described earlier, the APHAB—Aversiveness subscale was administered (Cox & Alexander, 1995). This is a six-item questionnaire that asks respondents about how often they are bothered by loud sounds. This measure was used as a predictor variable to determine whether the self-adjusted output, relative to the NAL target, was related to participants’ sensitivity to amplified sounds.

### Statistical Analyses

Repeated-measures analyses of variance (ANOVAs) and regression analyses were used to examine differences among, and relationships between, variables of interest. Greenhouse–Geisser adjustments were used, as needed, to correct for violations of sphericity (Greenhouse & Geisser, 1959). For significant effects or interactions, post hoc testing was completed using the Tukey (1949) Honest Significant Difference test.

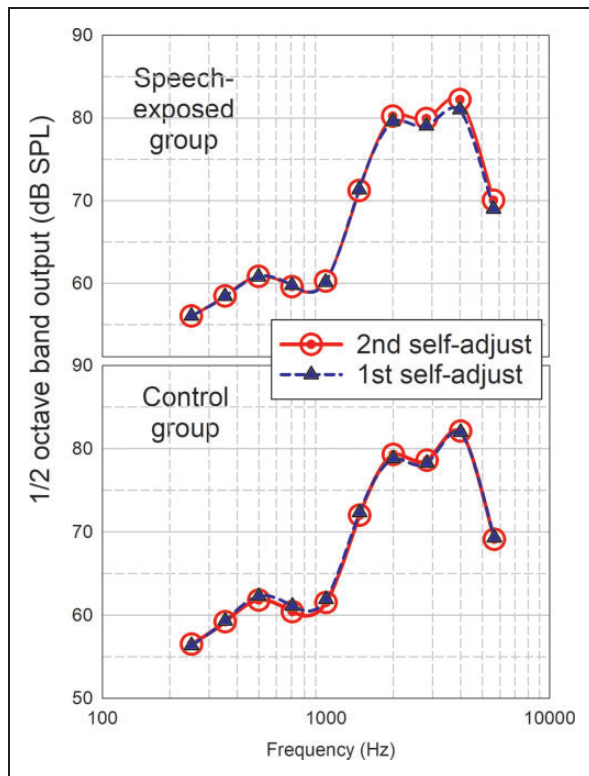
## Results

### Exploration

All participants completed the self-adjustment task. The amount of exploration varied among participants, but none accepted the starting response without exploration. The average range of exploration was 17 dB for overall amplitude, 9 dB for low-frequency cut (at 354 Hz) and 9 dB for high-frequency boost (at 3 kHz).

### User-Adjusted Output Spectra

**Effect of Exposure to Speech-Recognition Tests.** The mean real-ear half-octave output spectra for the first and second adjustments, collapsed across the quiet and the noise adjustment conditions, are shown in Figure 5. The upper panel shows data for the group exposed to speech-perception tests between the two adjustments. The lower panel shows data for the control group completing a nonauditory task between the two adjustments.

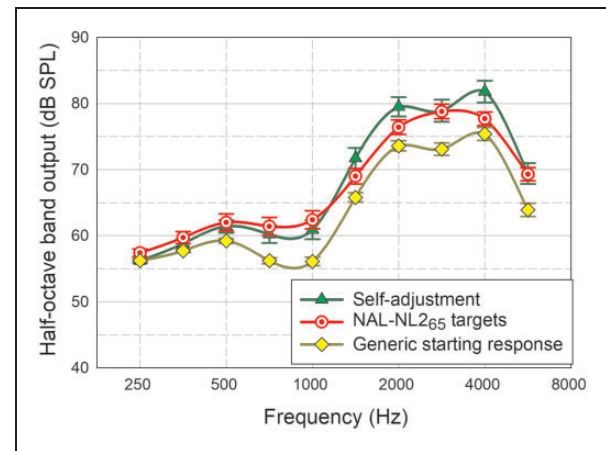


**Figure 5.** Group-Mean Real-Ear Output Levels as Functions of Frequency for the First and Second Self-Adjustments. Data for the exposure and control group are shown on the top and bottom, respectively.

Mean differences between the first and second adjustments were small (less than 2 dB) at all frequencies and for both participant groups.

To examine differences between and within groups in more detail, a mixed repeated-measures ANOVA was completed. Repeated measures were replication (first and second self-adjustment), noise condition (with and without babble), and frequency (10 levels: 250–6000 Hz in half-octave steps). Exposure group (exposure vs. control) was the between-subject factor. The interaction between group, replication, and frequency, already illustrated in Figure 5, was not significant,  $F(2.8, 61.4) = 0.44, p = .93, \eta_p^2 = 0.02$ . In addition, there was no significant main effect of group and no significant two-way interactions between group and any other factor. These data do not support the conclusion that administration of a formal speech-perception test after the first self-adjustment affected group-mean outcome after the second.

**Effect of Multitalker Babble.** In the analysis just described, the main effect of noise failed to reach the .05 level of significance. There was a significant interaction between noise condition and frequency, but with Greenhouse–Geisser correction for degrees of freedom, even this



**Figure 6.** Group-Mean Real-Ear Outputs ( $\pm 1$  Standard Error) as a Function of Frequency for the Generic Starting Response, Self-Adjusted Response (Averaged Over Two Adjustments in Noise and Two in Quiet), and NAL-NL2 Targets. NAL-NL2 = National Acoustics Laboratories prescription for Non-Linear hearing aids.

interaction fell above the 5% level of significance,  $F(3.9, 86.0) = 2.19, p = .08$ . There were no other significant interactions involving noise.

These findings do not support the conclusion that the presence of multitalker babble at an SNR of +6 dB significantly affected the group-mean self-adjusted output, averaged across two self-adjustments.

**User-Selected and NAL-NL2 Target Output.** Because of the absence of any significant group or noise effect, data were collapsed across groups, replication, and noise condition for subsequent analyses. Group-mean output responses for the generic starting condition and the average of the four self-adjusted conditions (two in quiet, two in noise) are shown in Figure 6 together with the group-mean NAL-NL2 targets (the NAL-NL2 targets are included for comparison only, not to imply that they were an intended goal for these self-fittings). A repeated-measures ANOVA was completed using setting (generic starting response, user-adjusted response, NAL-NL2 target) and frequency as within-subject factors. After Greenhouse–Geisser correction for nonsphericity, there was a significant interaction between setting and frequency,  $F(4, 91) = 8.36, p < .0001, \eta_p^2 = 0.27$ . In Tukey post hoc testing, the self-adjusted output was significantly higher than the starting output at all frequencies above 500 Hz. Significant differences between self-adjusted and prescribed output were found at 1.5 kHz ( $p = .05$ ), 2 kHz ( $p = .005$ ), and 4 kHz ( $p = .00003$ ). These findings support the conclusion that self-adjusted output was higher than the starting output and, in the higher frequencies, higher than the NAL-NL2 prescription.

### Individual Differences

**High-Frequency Versus Low-Frequency Output.** The group-mean spectra were reproducible, with no significant change between the first and second self-adjustments. Individual self-adjustments, however, varied widely. To explore these individual differences, separate low- and high-frequency outputs were obtained from the real-ear output spectra by energy summation from 500 to 1000 Hz and from 2 to 4 kHz. These frequency bands were selected on the basis of the configuration of the output spectra shown in Figures 5 and 6. The equation for energy summation was as follows:

$$y = 10 * \log_{10} \sum 10^{(x/10)} \quad (1)$$

where  $y$  is the summed level in dB and  $x$  is a single half-octave level in dB, obtained from the Verifit real-ear measure for a speech input of 65 dB SPL.

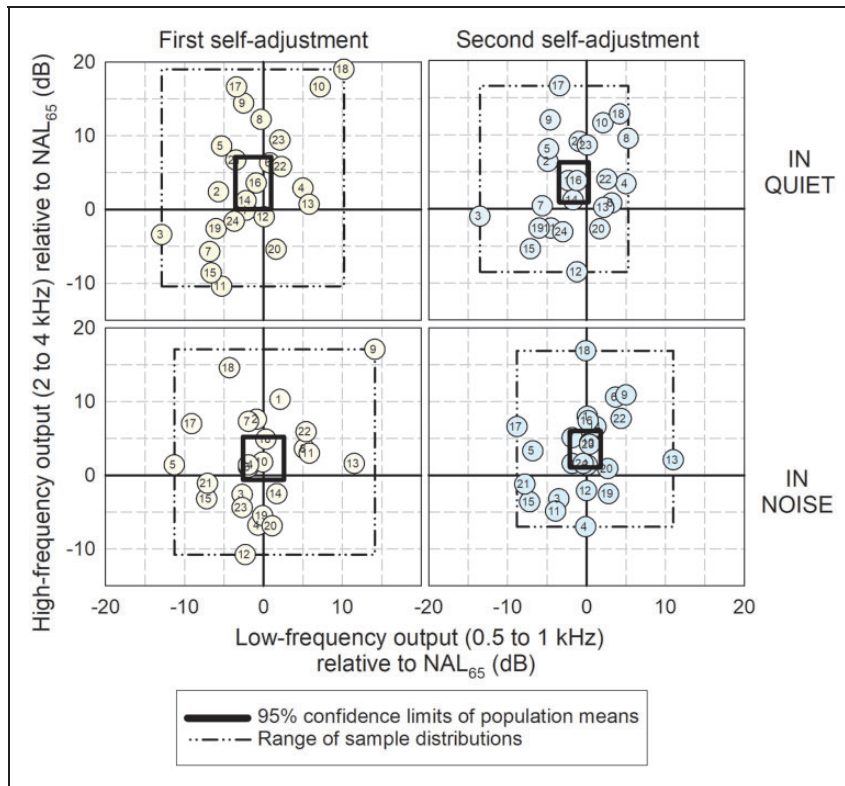
Figure 7 shows plots of high-frequency versus low-frequency output, relative to individual NAL-NL2 targets. First self-adjustments are shown on the left and second self-adjustments on the right. Adjustments in quiet are shown at the top and, in noise, at the bottom. Also shown are the 95% confidence levels for the population

means and the maxima and minima for the sample distributions. Participant number is shown within the symbols.

The sample distributions reveal a wide range of individual deviation from the NAL-NL2 prescription. After the second self-adjustment, for example, the low-frequency adjustments covered a range of 19 dB, from 14 dB below prescription to 5 dB above it. The high-frequency adjustments covered a range of 24 dB, from 7 dB below prescription to 17 dB above it. Similar ranges were found for the second self-adjustment in noise.

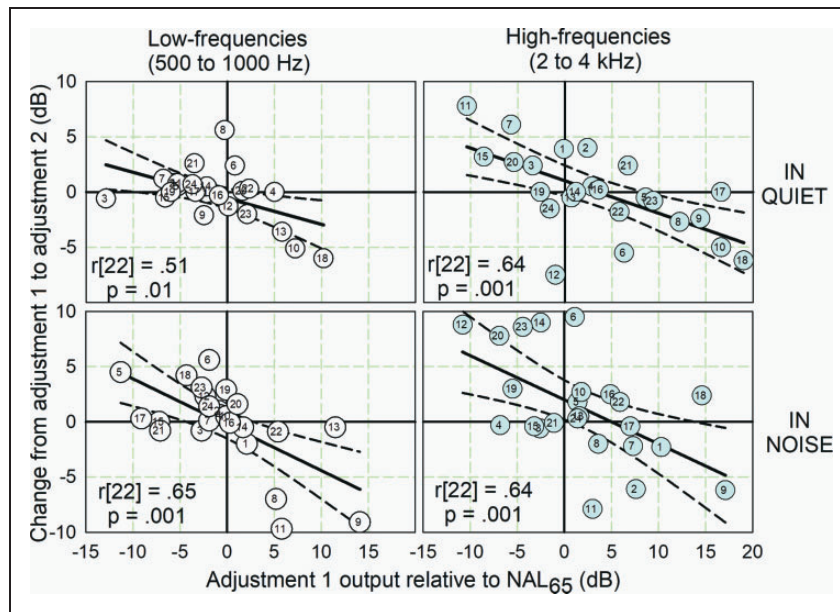
The ranges of the sample distributions in Figure 7 are less at the second self-adjustment than at the first, suggesting the possibility of changes toward prescription with repeated adjustment.

To explore this possibility in more detail, the change between the two adjustments was examined as a function of output at the first adjustment. The results are shown in Figure 8. Left and right panels show low- and high-frequency output. Upper and lower panels show adjustment in quiet and in noise. Linear regression functions with 95% confidence limits are shown in each panel. All four coefficients of correlation are significantly different from zero at either at the .01 level (low-frequency adjustments in quiet) or at the .001 level (all other adjustments). The confidence limits of the regression functions support

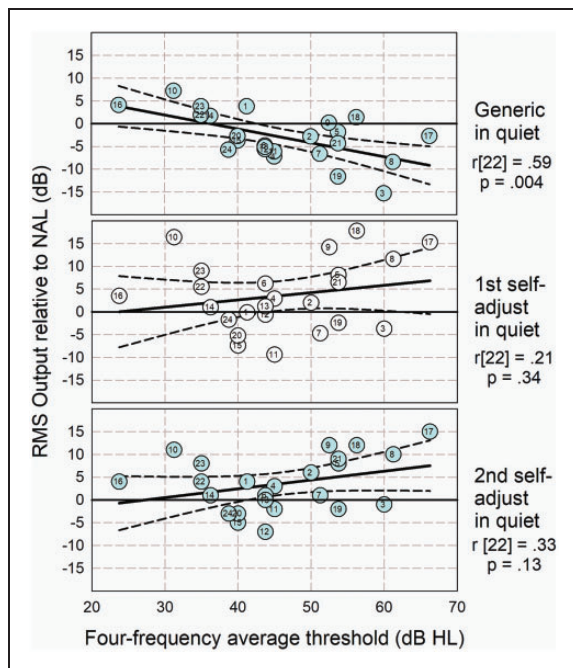


**Figure 7.** Distributions of High-Frequency Versus Low-Frequency Output for 24 Participants After Two Self-Adjustments in Quiet and Two in Noise. NAL = National Acoustics Laboratories.





**Figure 8.** Change of Low- and High-Frequency Output Between Adjustments as a Function of Value at the First Adjustment. NAL = National Acoustics Laboratories.



**Figure 9.** Speech Output in dB Relative to the NAL-NL2 Prescription, as a Function of Four-Frequency Average Threshold, for the Generic Starting Condition, and the First and Second Self-Adjustments. Lines show linear regression functions with 95% confidence limits. NAL = National Acoustics Laboratories; RMS = root mean square.

the conclusion that first self-adjustments well above prescription resulted in significant decreases at the second, while first self-adjustments well below prescription result in significant increase at the second. The predicted

population-mean changes, however, are small, amounting to 5 dB or less at the extremes.

*Effect of Hearing Loss on Differences From NAL Target.* Figure 9 shows, for the quiet condition only, the RMS output in dB relative to the NAL-NL2 target for each individual. These outputs are for a 65 dB SPL speech input and are shown as functions of four-frequency average pure-tone threshold (average of 0.5, 1, 2, and 4 kHz) in the test ear. The top, middle, and bottom panels show data for the generic starting condition, the first self-adjustment, and the second self-adjustment, respectively. Linear regression functions with 95% confidence limits are shown in each panel.

The significant correlation between hearing loss and the extent to which a fixed generic response falls below the prescriptive target is entirely predictable. There is no evidence from these data, however, to indicate that the self-adjusted deviation from the NAL-NL2 prescription, when measured in terms of overall RMS level, depends on degree of hearing loss.

*Repeatability of Self-Adjustment.* The repeatability of preferred self-adjusted RMS output in quiet was good. The coefficient of linear correlation between the first and second RMS outputs was .93. Group-mean difference was 0.2 dB with a standard deviation of 3.5 dB. Twenty-one of the 24 participants changed by 5 dB or less between the first and second self-adjustment. Repeatability in noise was poorer—with a correlation of .80. Group-mean difference was 1 dB with a standard deviation of 4.9 dB. Still, 16 of the 24 participants

changed by 5 dB or less between the first and second self-adjustment in noise.

### Speech Intelligibility Index

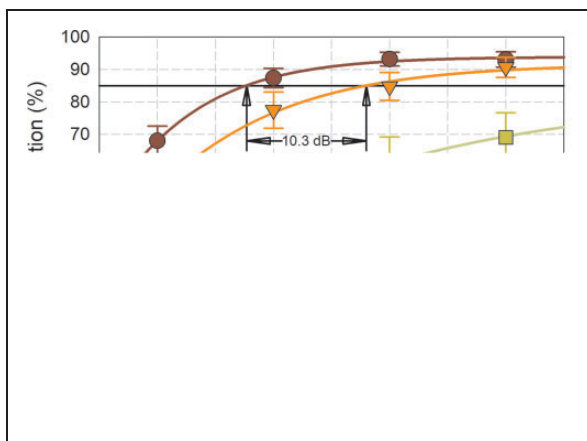
Mean SII in quiet was 0.66 under the generic condition, 0.76 averaged across the two self-adjustments, and 0.78 for the NAL-NL2 targets. An acceptable criterion of 0.6 was taken from the previous study, in which this value corresponded with 96% recognition of words in short sentences (Mackersie et al., 2019). Under the generic condition, 15 (63%) of the participants in this study met this criterion. After the first self-adjustment, 22 of the participants, (92%) did so and, after the second, 21 (88%). Had everyone adjusted to the NAL prescription, all would have met the 0.6 criterion of acceptable audibility used here.

### Speech Perception

**Phoneme-in-Word Recognition.** Mean phoneme recognition scores in quiet are shown as a function of speech level in Figure 10 for three conditions (unaided, generic starting response, and second self-adjustment). The horizontal line shows a lowest acceptable criterion of 85%; this criterion has been shown to correspond to 95% recognition of short sentences by normal-hearing listeners (Boothroyd & Nittrouer, 1988). The curves show least-squares cubed exponential fits to the mean data using the following equation:

$$y = a \times \left(1 - e^{-(x-b)/c}\right)^3 \quad (2)$$

where  $y$  is the percentage recognition score,  $a$  is the asymptotic (maximum) score in percentage,  $e$  is



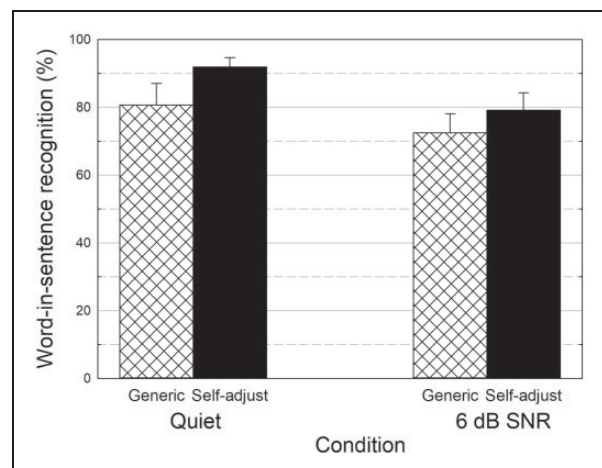
**Figure 10.** Group-Mean Phoneme Recognition ( $\pm 1$  Standard Error) as a Function of Speech Level for Three Listening Conditions. Curves are least-squares fits to Equation 2. RMS = root mean square.

the base of natural logarithms,  $x$  is speech level in dB SPL, and  $b$  is the speech level in dB SPL at which recognition falls to zero.

A repeated-measures ANOVA of arcsine-transformed recognition scores, using condition and level as within-subject factors, showed a main effect of condition,  $F(1.2, 27.6) = 27.17$ ,  $p < .0001$ ,  $\eta_p^2 = 0.54$ , and level,  $F(1.9, 42.5) = 80.12$ ,  $p < .0001$ ,  $\eta_p^2 = 0.78$ . There was a significant interaction between condition and level,  $F(3.5, 79.6) = 2.15$ ,  $p = .05$ ,  $\eta_p^2 = 0.09$ . Post hoc tests indicated significant differences ( $p < .05$ ) among all three conditions at 45, 55, and 65 dB SPL, but not between the starting condition and the second self-adjustment at 75 dB. Group-mean scores reached the 85% criterion for a speech input of 53.7 dB SPL after self-adjustment, but not until 64.0 dB SPL for the starting condition, a difference of 10.3 dB.

These data support the conclusion that self-adjustment from the generic starting condition resulted in significantly improved speech perception.

**Word-in-Sentence Recognition.** Mean scores for the CUNY sentences presented at 60 dB SPL are shown in Figure 11 for the generic and second self-adjusted settings in both quiet and noise. Scores obtained with the self-adjusted setting were generally higher than those with the generic setting, especially in quiet. A repeated-measures ANOVA, using condition (quiet, noise) and setting (generic, user-adjusted) indicated a main effect of both condition,  $F(1, 23) = 30.36$ ,  $p < .0001$ ,  $\eta_p^2 = 0.57$ , and setting,  $F(1, 23) = 9.18$ ,  $p = .006$ ,  $\eta_p^2 = 0.29$ . These data demonstrate positive effects of self-adjustment and negative effects of noise, even in a task that enables maximal use of sentence context to compensate for reduced audibility.

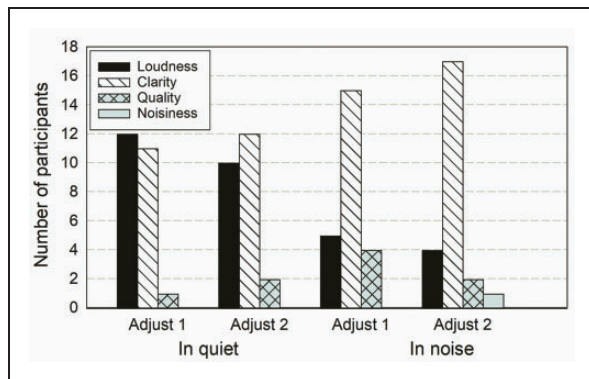


**Figure 11.** Group-Mean Effects of Self-Adjustment and Noise on Word-in-Sentence Scores. Error bars show one standard error. SNR = signal-to-noise ratio.

**Table 2.** Mean Self-Reported Ratings of the Importance of Four Subjective Adjustment Criteria on a 5-Point Scale (1 = Not at All; 5 = Very).

Noise condition	Adjustment number	Criterion			
		Loudness	Clarity	Quality	Noisiness
Quiet	First	4.5	4.4	4.0	2.1
	Second	4.4	4.3	4.1	2.0
6dB SNR	First	4.3	4.5	4.1	3.0
	Second	4.3	4.4	4.0	3.3

Note. SNR = signal-to-noise ratio.



**Figure 12.** Number of Times Each of Four Adjustment Criteria Was Selected as Being Most Important by the 24 Participants.

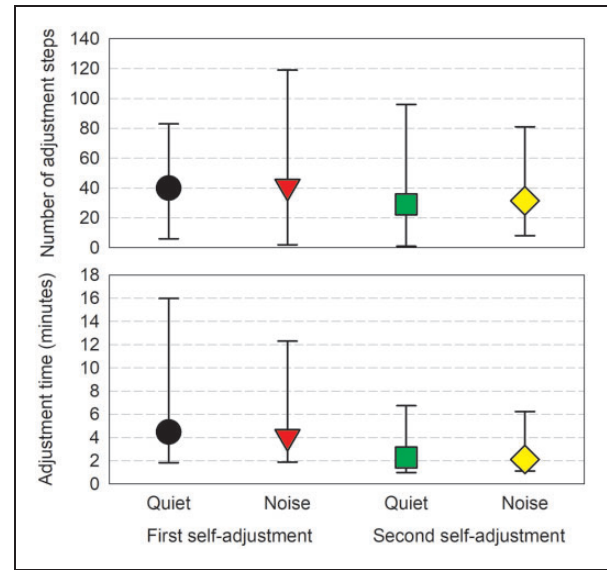
*Listener Criteria for Self-Adjustment*

Table 2 shows the group mean ratings of the importance of four subjective criteria when making self-adjustments. Loudness, clarity, and quality all received high ratings in both quiet and noise. The rating of noisiness was much lower (but not absent) in quiet but increased by an average of 1 point when noise was actually present.

The number of times each criterion was selected as being *the most important* is shown in Figure 12. In quiet, the differences between clarity and loudness were small. When adjusting in noise, however, clarity was the dominant criterion. The effect of noise on the relative importance of loudness and clarity was significant ( $\chi^2 = 6.8$ ;  $p = .01$ ). Quality and noisiness were seldom selected as the most important criterion.

*Number of Steps and Time*

Figure 13 shows the average and range for the adjustment time and number of adjustment steps at the first and second adjustments in quiet and in noise. The first self-adjustment took an average of 41 steps in quiet and 40 steps in noise. The number of steps at the second self-adjustment fell by around 25% to 29 in quiet and 31 in noise. A repeated-measures ANOVA of the number of steps using replication and noise as within-subjects



**Figure 13.** Average and Range (Error Bars) of Steps Taken (Upper Panel) and Time Taken (Lower Panel) During Completion of a Single Self-Adjustment.

factors indicated a main effect of replication,  $F(1, 23) = 7.8$ ,  $p = .01$ . There was no main effect of noise and no significant interaction.

The first self-adjustment took an average of 4 min, 28 s in quiet and 4 min, 3 s in noise. These times fell by around 50% to 2 min, 17 s and 2 min, 6 s for the second self-adjustment. These data were log transformed to correct the positive skew and analyzed using an ANOVA with replication and noise condition as factors. There was a significant main effect of replication,  $F(1, 23) = 8.0$ ,  $p = .009$ , but no main effect of noise and no significant interaction. At the second adjustment, no participant took more than 7 min to complete the process.

*Nonaudiological Predictors of Outcome*

Age, years of hearing-aid use, years of education, sex, Montreal Cognitive Assessment score, and sound aversiveness (APHAB) were examined in terms of their ability to predict overall RMS deviations from prescribed NAL-NL2 targets. None were found to explain significant amounts of variance among participants.

**Discussion**

A primary goal of this study was to determine the benefit of including a formal speech-perception test as a component of the self-adjustment process. The previous study (Boothroyd & Mackersie, 2017; Mackersie et al, 2019) showed a group-mean change of adjusted output by hearing-aid users after (though not necessarily because of) a formal speech-perception test, taken while using an initial adjustment. No such effect was

observed in this study. There was, however, an important difference between this and the previous study, namely, the use of real-time processing of sound-field microphone input. The open microphone allowed participants in both groups to hear the researcher's instructions and their own speech before and after the first adjustment. It is possible that this experience was enough to eliminate group mean changes in a second adjustment, with or without administration of a formal speech-perception test. The presence, content, and extent of this informal exposure were not planned as independent variables in this study. Consequently, there are no data to support conclusions about their relative contributions. Note, however, that the speech of the audiologist and the client often provide the sole listening experience on which initial fine-tuning is based in current clinical practice.

Even though the formal speech perception test had no significant effect on group-mean adjustments, there was statistical evidence of change in the direction of the NAL-NL2 targets between the first and second adjustments, at least for participants with the greatest deviation after the first adjustment. This evidence points to the potential value of at least two self-adjustments during an initial self-fitting.

The group-mean self-adjustments from a generic starting response were either not significantly different from or exceeded (by up to 5 dB) group-mean NAL-NL2 prescription. Individual adjustments, however, varied from individualize prescriptions by varying amounts. Similar findings were reported by Nelson et al. (2018). The maximum deviations found here and shown in Figure 8 (14 for low frequencies and 21 dB for high frequencies) are considerably lower than the 24 and 38 dB reported by Nelson et al. There were, however, marked differences in equipment and procedure for the two studies. Nevertheless, the findings of variability in the two studies underline the fact that the NAL-NL2 prescription is not intended to be, or claimed to be, the ideal response for every individual with a given audiometric configuration—only for the average of many individuals with that configuration. Even for a single individual, there can be a range of settings providing an acceptable compromise among loudness, comfort, sound quality, and intelligibility. Within-subject variation of self-adjusted output could well represent different choices of placement within an acceptable range. Such choices could also have affected individual changes from Adjustment 1 to Adjustment 2 in this study.

While group-mean outputs and spectra were at or close to the group-mean NAL-NL2 prescription, it is important to note that the comparisons reported here were based only on a speech input of 65 dB SPL. The compression ratio of 1.4:1 used in this study was lower than the value of 2:1 or more that would have been

prescribed. As a result, the gain and output for a 45 dB SPL input were lower than prescription—perhaps accounting for the rapid fall of group-mean phoneme recognition for inputs below 60 dB SPL shown in Figure 10.

Some participants reported difficulty hearing the effects of changes of Fullness (i.e., of low-frequency output). One possible reason relates to the slope of the skirts of the low-frequency band-pass filters. In the previous study, these slopes were deliberately made very steep and were consistent across frequency. With real-time processing, however, time constraints limit the steepness of FIR filter skirts at lower frequencies. As a result, the maximum attenuation in a low-frequency band is limited to that in the skirt of the filter with the next higher frequency.

Another reason for difficulty hearing the effect of changes of low-frequency output is leakage of sound past, or through, the dome used with the ear-canal receivers (Balling et al., 2019). This leakage allows low frequencies from the sound field to enter the ear canal. At the same time, it allows low frequencies from the MHA to escape. Once the amplitude of the entering sound exceeds that of the MHA output, difficulty hearing the effects of changes in the latter is inevitable.

A second problem with the real-time system was instability resulting from acoustic feedback. Although the amplification software included acoustic feedback management, this became ineffective if a participant pushed high-frequency gain beyond a certain point. Four of the five participants who experienced feedback issues during the self-fitting procedure had a high-frequency average hearing loss in excess of 60 dB. Not only did these thresholds call for high gain in the higher frequencies, but these listeners also preferred an overall output level that was more than 7 dB above that prescribed by NAL-NL2. Subsequent iterations of the Goldilocks software platform have addressed this problem by placing a researcher-adjustable limit on high-frequency gain.

There was no evidence that group-mean self-adjustments in multitalker babble at an SNR of +6 dB were different from those in quiet. Nelson et al. (2018) reported significant effects of noise on self-selected gain but, at an SNR of +5 dB, the effect was small. The finding of no significant effect at an SNR of +6 dB in this study does not mean that audibility was unaffected. The effects of noise were clearly demonstrated in terms of word-in-sentence recognition.

The mean time taken to complete the first adjustment (4 min, 15 s) was substantially longer than the mean time (1 min, 5 s) taken by experienced hearing-aid users in the previous study (Boothroyd & Mackersie, 2017). A possible explanation is the larger number of high-frequency steps available to participants resulting from the smaller

step size (3 dB step-size in this study compared with 5 dB in the earlier study). In addition, this study explicitly required the user to fully explore the limits of acceptability at both the high and low ends of the ranges for each of the three adjustment parameters. In the previous study, exploration was encouraged but not required.

Both this and the previous study showed reduction of steps and time between the first and second self-adjustments. This reduction could well reflect task-related learning. At the first adjustment, participants in this study spent an average of 6.3 s listening at each step. During the second adjustment, not only did they explore fewer steps but the listening time fell to 4.4 s per step.

The criteria selected as being most important were almost equally balanced between loudness and clarity for adjustments made in quiet. This may be interpreted as a balance between comfort and perceived, or estimated, intelligibility. The finding that participants selected clarity as being more important than loudness, when adjusting in noise, is consistent with the notion that participants were trying to improve perceived intelligibility when audibility was limited by noise and hearing threshold rather than by threshold alone. Note from Figure 13, however, that there was no evidence of an increase in time taken or number of steps explored when adjusting in noise, suggesting that the choice of most important criterion showed awareness of the effect of the noise on intelligibility rather than a change in strategy.

The SII data showed acceptable speech-weighted audibility for just over half of these participants when listening with the generic starting condition to speech with a level of 65 dB SPL. This speaks quite well for the viability of a “one-size-fits-all” version of a direct-to-consumer hearing aid, but the fact that around 90% reached this criterion after self-adjustment supports the value of self-adjustment. Selecting a single criterion for an acceptable value of SII is, of course, difficult. Data from listeners with normal hearing strongly suggest an SII criterion for acceptability that is below 0.6. In fact, the results of a study by Sherbecoe and Studebaker (2002) suggest that a criterion of 0.6 should provide word-in-sentence recognition scores of around 98% in normally hearing listeners. A higher criterion for listeners with hearing loss acknowledges the deficits of spectral and temporal resolution accompanying cochlear damage together speech-perception difficulties associated with aging. Nevertheless, without more research evidence, the selection of 0.6 as an acceptable SII criterion for listeners with mild-to-moderate hearing loss must remain somewhat arbitrary.

Much of the research on self-fitting is based on the assumption that the listener’s self-adjustment should start from a threshold-based prescription. There is, indeed, evidence that the starting point for self-adjustment can affect the end point (Dreschler et al.,

2008; Keidser et al., 2008; Mueller et al., 2008). A prescriptive starting point, before fine-tuning, is also in keeping with the current standard of clinical audiological practice. But this and the previous study (Mackersie et al., 2019) found that the Goldilocks procedure, starting from a generic response, gave group-mean outputs that were barely distinguishable from NAL-NL2 prescriptions. This finding suggests that a threshold-based starting response may not be necessary for self-fitting of direct-to-consumer hearing aids. Note, however, that the starting response used in these studies was, in fact, based on an NAL-NL2 prescription for a generic threshold configuration that was close to that of several of these participants. Whether an individual threshold-based starting response would provide results closer to the NAL-NL2 prescription, for individuals whose audiograms are very different from the generic audiogram used here, has yet to be determined.

This and our previous study were restricted to self-adjustment of level and spectrum. Efficacy of, and candidacy for, self-adjustment of such things as compression characteristics, maximum output, noise management, and directionality has yet to be explored in detail. Also in need of further study is self-adjustment of binaural amplification.

Although this study did not use binaural amplification, the nontest ears of participants with thresholds of 40 dB or better were occluded. Nevertheless, higher sound-field inputs used for the Computer-Assisted Speech-Perception Assessment testing (75 dB SPL) might have been partially audible via the unaided ear for some participants. Data showing the unaided phoneme recognition scores to be considerably poorer than those obtained with monaural amplification (Figure 10), however, do suggest minimal contribution from the unaided ear under the amplified conditions.

The research reported here used a specific “explore-and-select” approach to user self-adjustment. There are alternative strategies involving such things as paired comparison and machine learning. Some of these strategies extend to other factors such as the criteria used by listeners and the properties of the sound input during self-adjustment (e.g., Jensen et al., 2019). Empirical comparisons of methods for self-fitting and postfitting self-readjustment are clearly needed.

## Conclusions

1. The administration of a formal speech-perception test, after a first self-adjustment, did not have a significant effect on self-adjusted real-ear output at the second.
2. The presence of multitalker babble at an SNR of +6 dB did not have a significant effect on self-adjusted real-ear output. It did, however, affect

subjective opinions on the most important subjective criterion guiding adjustment.

3. Although all self-adjustments began with the same generic starting response, rather than individual NAL-NL2 prescriptions, the group-mean self-adjusted real-ear output was not significantly different from the group-mean prescription at most frequencies and exceeded it (by up to 5 dB) at some.
4. Individual self-adjusted outputs were both higher and lower than prescription in both low and high frequencies. Participants with the largest deviations from prescription at the first self-adjustment made small but statistically significant changes in the direction of prescription at the second, leaving the group means essentially unchanged.
5. The number of adjustments made, and the time taken, fell significantly between the two self-adjustments, suggesting task-related learning.

These findings continue to support the efficacy of hearing-aid self-fitting and postfitting readjustment by adults with mild-to-moderate hearing loss.

### Acknowledgments

The assistance of Robert Novak, Brittany Thrall, Shannon Doolittle, and Jennifer Retana at San Diego State University and Mark Stambaugh, Adit Bharadwaj, Swaroop Gadiyaram, and Krishna Chaitanya Vastare at the University of California, San Diego is gratefully acknowledged.

### Data Accessibility Statement

Data will be made available by request at the end of the project per the approved National Institutes of Health resource sharing plan for the grant.



### Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research reported in this publication was supported by the National Institute on Deafness and other Communication Disorders of the National Institutes of Health (Grant Number R33DC015046). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

### ORCID iDs

Carol L. Mackersie  <https://orcid.org/0000-0003-4789-5588>  
Arthur Boothroyd  <https://orcid.org/0000-0002-3103-7290>

### References

- Abrams, H. (2017). Hearing loss and associated comorbidities: What do we know? *Hearing Review*, 24(12), 32–35.
- Almufarrij, I., Munro, K. J., Dawes, P., Stone, M. A., & Dillon, H. (2019). Direct-to-consumer hearing devices: Capabilities, costs, and cosmetics. *Trends in Hearing*, 23, 2331216519858301. <https://doi.org/10.1177/2331216519858301>
- ANSI (2014). *Specification of hearing aid characteristics (ANSI S3.22-2014)* New York, NY, American National Standards Institute.
- ANSI (1997). *Methods for the calculation the speech intelligibility index (ANSI S3.5 -1997 R2012)* New York, NY, American National Standards Institute.
- Balling, L. W., Jensen, N. S., Caporali, S., Cubick, J., & Switalski, W. (2019). Challenges of instant-fit ear tips: What happens at the eardrum? *Hearing Review*, 26(12), 12–15.
- Besser, J., Stropahl, M., Urry, E., & Launer, S. (2018). Comorbidities of hearing loss and the implications of multimorbidity for audiological care. *Hearing Research*, 369, 3–14. <https://doi.org/10.1016/j.heares.2018.06.008>
- Boothroyd, A. (1968). Developments in speech audiometry. *British Journal of Audiology*, 2, 3–10. <https://doi.org/10.3109/05384916809074343>
- Boothroyd, A. (2008). The performance/intensity function: An underused resource. *Ear and Hearing*, 29(4), 479–491. <https://doi.org/10.1097/AUD.0b013e318174f067>
- Boothroyd, A., Hnath-Chisolm, T., Hanin, L., & Kishon-Rabin, L. (1988). Voice fundamental frequency as an auditory supplement to the speechreading of sentences. *Ear and Hearing*, 9(6), 306–312. <https://doi.org/10.1097/00003446-198812000-00006>
- Boothroyd, A., & Mackersie, C. L. (2017). A “Goldilocks” approach to hearing-aid self-fitting: User interactions. *American Journal of Audiology*, 26, 430–435. [https://doi.org/10.1044/2017\\_aja-16-0125](https://doi.org/10.1044/2017_aja-16-0125)
- Boothroyd, A., & Nitttrouer, S. (1988). Mathematical treatment of context effects in phoneme and word recognition. *Journal of the Acoustical Society of America*, 84, 101–114. <https://doi.org/10.1121/1.396976>
- Brodie, A., & Ray, J. (2018). Generic quality of life in persons with hearing loss: A review of the recent literature. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 39(9), 1074–1078. <https://doi.org/10.1097/MAO.0000000000001959>
- Brody, L., Wu, Y.-H., & Stangl, E. (2018). A comparison of personal sound amplification products and hearing aids in ecologically relevant test environments. *American Journal of Audiology*, 27(4), 581–593. [https://doi.org/10.1044/2018\\_AJA-18-0027](https://doi.org/10.1044/2018_AJA-18-0027)
- Ciletti, L., & Flamme, G. A. (2008). Prevalence of hearing impairment by gender and audiometric configuration: Results from the National Health and Nutrition Examination Survey (1999–2004) and the Keokuk County Rural Health Study (1994–1998). *Journal of the American Academy of Audiology*, 19, 672–682. <https://doi.org/10.3766/jaaa.19.9.3>

- Cox, R., & Alexander, G. (1995). Abbreviated profile of hearing aid benefit. *Ear and Hearing, 16*, 176–183. <https://doi.org/10.1097/00003446-199504000-00005>
- Dreschler, W. A., Keidser, G., Convery, E., & Dillon, H. (2008). Client-based adjustments of hearing aid gain: The effect of different control configurations. *Ear and Hearing, 29*(2), 214–227. <https://doi.org/10.1097/AUD.0b013e31816453a6>
- Garudadri, H., Boothroyd, A., Lee, C. H., Gadiyaram, S., Bell, J., Se ngupta, D., Hamilton, S., Vastare, K. C., Gupta, R., & Rao, B. D. (2017). *A realtime, open-source speech-processing platform for research in hearing loss compensation* [Paper presentation]. 51st Conference on Signals, Systems, and Computers, Asilomar, CA, United States. <https://doi.org/10.1109/acssc.2017.8335694>
- Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*, 95–112. <https://doi.org/10.1007/bf02289823>
- Hanin, L., Boothroyd, A., & Hnath-Chisolm, T. (1988). Tactile presentation of voice fundamental frequency as an aid to the speechreading of sentences. *Ear and Hearing, 9*(6), 335–341. <https://doi.org/10.1097/00003446-198812000-00010>
- Hogan, A., O'Loughlin, K., Miller, P., & Kendig, H. (2009). The health impact of a hearing disability on older people in Australia. *Journal of Aging and Health, 21*(8), 1098–1111. <https://doi.org/10.1177/0898264309347821>
- Humes, L. E., Rogers, S. E., Quigley, T. M., Main, A. K., Kinney, D. L., & Herring, C. (2017). The effects of service-delivery model and purchase price on hearing-aid outcomes in older adults: A randomized double-blind placebo-controlled clinical trial. *American Journal of Audiology, 26*(1), 53–79. [https://doi.org/10.1044/2017\\_AJA-16-0111](https://doi.org/10.1044/2017_AJA-16-0111)
- Jensen, N. S., Hau, O., Nielsen, J. B., Nielsen, T. B., & Legarth, S. V. (2019). Perceptual effects of adjusting hearing-aid gain by means of a machine-learning approach based on individual user preference. *Trends in Hearing, 23*, 2331216519847413. <https://doi.org/10.1177/2331216519847413>
- Keidser, G., & Convery, E. (2018). Outcomes with a self-fitting hearing aid. *Trends in Hearing, 22*, 2331216518768958. <https://doi.org/10.1177/2331216518768958>
- Keidser, G., Dillon, H., & Convery, E. (2008). The effect of the base line response on self-adjustments of hearing aid gain. *Journal of the Acoustical Society of America, 124*(3), 1668–1681. <https://doi.org/10.1121/1.2951500>
- Keidser, G., Dillon, H., Flax, M., Ching, T., & Brewer, S. (2011). The NAL-NL2 prescription procedure. *Audiology Research, 1*(1), e24–e24. <https://doi.org/10.4081/audiore.2011.e24>
- Knudsen, L. V., Oberg, M., Nielsen, C., Naylor, G., & Kramer, S. E. (2010). Factors influencing help seeking, hearing aid uptake, hearing aid use and satisfaction with hearing aids: A review of the literature. *Trends in Amplification, 14*(3), 127–154. <https://doi.org/10.1177/1084713810385712>
- Kochkin, S. (2010). MarkeTrak VIII: Utilization of PSAPs and direct-mail hearing aids by people with hearing impairment. *Hearing Review, 17*(12), 14–16.
- Lin, F. R., Thorpe, R., Gordon-Salant, S., & Ferrucci, L. (2011). Hearing loss prevalence and risk factors among older adults in the United States. *Journals of Gerontology Series A: Biological Sciences & Medical Sciences, 66A*(5), 582–590. <https://doi.org/10.1093/gerona/66a5>
- Mackersie, C. L., Boothroyd, A., & Lithgow, A. (2019). A “Goldilocks” approach to hearing aid self-fitting: Ear-canal output and Speech Intelligibility Index. *Ear and Hearing, 40*(1), 107–115. <https://doi.org/10.1097/AUD.0000000000000617>
- Mackersie, C. L., Boothroyd, A., & Minniear, D. (2001). Evaluation of the Computer-Assisted Speech Perception Assessment Test (CASPA). *Journal of the American Academy of Audiology, 12*(8), 390–396.
- Mueller, H. G., Hornsby, B. W. Y., & Weber, J. E. (2008). Using trainable hearing aids to examine real-world preferred gain. *Journal of the American Academy of Audiology, 19*(10), 758–773. <https://doi.org/10.3766/jaaa.19.10.4>
- Nasreddine, Z. S., Phillips, N. A., Bedirian, 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. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>
- Nelson, P. B., Perry, T. T., Gregan, M., & VanTasell, D. (2018). Self-adjusted amplification parameters produce large between-subject variability and preserve speech intelligibility. *Trends in Hearing, 22*, 2331216518798264. <https://doi.org/10.1177/2331216518798264>
- Over-the-Counter Hearing Aid Act of 2017, S.670. Amendment of Section 520 of the Federal Food, Drug, and Cosmetic Act (21 U.S.C. 360j). (2017).
- Perry, T. T., Nelson, P. B., & Van Tasell, D. J. (2019). Listener factors explain little variability in self-adjusted hearing aid gain. *Trends in Hearing, 23*, 2331216519837124. <https://doi.org/10.1177/2331216519837124>
- Powers, T. A., & Rogin, C. M. (2019). Marke Trak 10: Hearing aids in an era of disruption and DTC/OTC devices. *Hearing Review, 26*(8), 12–20.
- Punch, J. L., & Robb, R. (1992). Prescriptive hearing aid fitting by parameter adjustment and selection. *Journal of the American Academy of Audiology, 3*(2), 94–100.
- Punch, J. L., Robb, R., & Shovels, A. H. (1994). Aided listener preferences in laboratory versus real-world environments. *Ear and Hearing, 15*(1), 50–61. <https://doi.org/10.1097/00003446-199402000-00006>
- Reed, N. S., Betz, J., Kendig, N., Korczak, M., & Lin, F. R. (2017). Personal sound amplification products vs a conventional hearing aid for speech understanding in noise. *Journal of American Medical Association, 318*(1), 89–90. <https://doi.org/10.1001/jama.2017.6905>
- Sacco, G., Gonfrier, S., Teboul, B., Gahide, I., Prate, F., Demory-Zory, M., Turpin, J. M., Vuagnoux, C., Genovese, P., Schneider, S., & Guevara, N. (2016). Clinical evaluation of an over-the-counter hearing aid (TEO First®) in elderly patients suffering of mild to moderate hearing loss. *BMC Geriatrics, 17*, 1–7. <https://doi.org/10.1186/s12877-016-0304-4>
- Sherbecoe R.L. and Studebaker G.A. (2002) Audibility-index functions for the connected speech test. *Ear and Hearing,*

- 23, 385–398. <https://doi.org/10.1097/00003446-200210000-0000>
- Souza, P. E., Arehart, K. H., Shen, J., Anderson, M., & Kates, J. M. (2015). Working memory and intelligibility of hearing-aid processed speech. *Frontiers in Psychology, 6*, 526–526. <https://doi.org/10.3389/fpsyg.2015.00526>
- Tahden, M. A. S., Gieseler, A., Meis, M., Wagener, K. C., & Colonius, H. (2018). What keeps older adults with hearing impairment from adopting hearing aids? *Trends in Hearing, 22*, 2331216518809737. <https://doi.org/10.1177/2331216518809737>
- Tukey, J. (1949). Comparing individual means in the analysis of variance. *Biometrics, 5*(2), 99–114. <https://doi.org/10.2307/3001913>