

# Short-Term and Long-Term Hearing Aid Benefit and User Satisfaction: A Comparison between Two Fitting Protocols

Lu-Feng Shi\*  
Karen A. Doherty\*  
Tammy M. Kordas\*  
Joseph T. Pellegrino\*

## Abstract

Currently published hearing aid fitting protocols recommend speech-in-noise testing and loudness measures, but it remains unclear how these measures affect hearing aid benefit and user satisfaction. This study compared two protocols in their effects on benefit and satisfaction. Protocol A included an electroacoustic analysis, real-ear measures, and hearing aid adjustments based on users' comments. Protocol B included all of Protocol A and a speech-in-noise test, loudness discomfort levels, and aided loudness. Thirty-two participants completed the Abbreviated Profile of Hearing Aid Benefit (APHAB) and the Satisfaction with Amplification in Daily Life (SADL) at 45 days and three months post-initial fitting. Fewer hearing aid adjustments were made to the hearing aids for participants fitted with Protocol B than participants fitted with Protocol A, but final gains were similar for both groups. Although similar APHAB scores were obtained for both protocols, SADL scores decreased between 45 days and three months for Protocol A.

**Key Words:** Benefit, fitting protocol, hearing aids, satisfaction

**Abbreviations:** ANSI = American National Standards Institute; APHAB = Abbreviated Profile of Hearing Aid Benefit; ASHA = American Speech-Language-Hearing Association; AV = Aversiveness of sound (APHAB subscale); BN = Background Noise (APHAB subscale); DSL[i/o] = Desired Sensation Level [input/output]; LDL = loudness discomfort level; NF = Negative Feature (SADL subscale); PI = Personal Image (SADL subscale); QuickSIN = Quick Speech-in-Noise test; REAR = real-ear aided response; REUR = real-ear unaided response; SADL = Satisfaction with Amplification in Daily Life; SC = Service and Cost (SADL subscale); SNR = signal-to-noise ratio; SRT = speech recognition threshold

## Sumario

Los protocolos de amplificación de auxiliares auditivo actualmente publicados recomiendan pruebas de lenguaje en ruido y mediciones de apreciación subjetiva de la intensidad (sonoridad), pero no está claro cómo estas mediciones afectan el beneficio de un auxiliar auditivo y la satisfacción del usuario. El estudio comparó dos protocolos en cuanto a sus efectos sobre beneficio y satisfacción. El Protocolo A incluyó un análisis electroacústico, mediciones de oído real y ajuste en el auxiliar auditivo basados en los comentarios del usuario. El Protocolo B incluyó todas las pruebas del Protocolo A, además de una prueba de audición en ruido, de niveles de molestia en la apreciación subjetiva de la

---

\*Gebbie Hearing Clinic, Department of Communication Sciences and Disorders, Syracuse University, Syracuse, New York

Lu-Feng Shi, Department of Communication Sciences and Disorders, Long Island University–Brooklyn Campus, Brooklyn, New York 11201; Phone: 718-780-4169; Fax: 718-780-4007; E-mail: lu.shi@liu.edu

Portions of this work were presented at the 2004 ASHA Convention in Philadelphia, Pennsylvania, and the 2005 ASHA Convention in San Diego, California.

intensidad y de sonoridad amplificadas. Treinta y dos participantes completaron el Perfil Abreviado de Beneficio del Auxiliar Auditivo (APHAB) y la prueba de Satisfacción con la Amplificación en la Vida Diaria (SADL) a los 45 días y a los tres meses de la adaptación inicial. Tuvieron que hacerse menos ajustes en el audífono en los auxiliares auditivos de participantes adaptados con el Protocolo B, que en los participantes adaptados con el Protocolo A, pero las ganancias finales fueron similares en ambos grupos. Aunque se obtuvieron puntajes APHAB similares en ambos protocolos, los puntajes SADL disminuyeron entre los 45 días y los tres meses para el Protocolo A.

**Palabras Clave:** Beneficio, protocolo de adaptación, auxiliares auditivos, satisfacción

**Abreviaturas:** ANSI = Instituto Nacional Americano de Estándares; APHAB = Perfil Abreviado de Beneficio con el Auxiliar Auditivo; ASHA = Asociación Americana de Audición y Lenguaje; AV = Aversión al sonido (sub-escala APHAB); BN = Ruido de Fondo (sub-escala APHAB); DSL[i/o] = Nivel de Sensación Deseado (ingreso/salida); LDL = nivel molesto de sonoridad; NF = Rasgo Negativo (sub-escala SADL); PI = Imagen Personal (sub-escala SADL); QuickSIN = Prueba Rápida de Lenguaje en Ruido; REAR = respuesta amplificadas en oído real; REUR = respuesta no amplificadas en oído real; SADL = Satisfacción con la amplificación en la vida real; SC = Servicio y Costo (sub-escala SADL); SNR = Tasa señal/ruido; SRT = umbral de reconocimiento del lenguaje

Many guidelines have been suggested in the fitting of hearing aids, but no single protocol is universally accepted. Despite differences, published guidelines share some of the same recommendations (e.g., ASHA [American Speech-Language-Hearing Association], 1998; Kochkin, 2003; Mueller, 2003; Washington University School of Medicine, 2004). All protocols recommend obtaining the 2-cc coupler response of a hearing aid to compare it with the ANSI (American National Standards Institute) standards (1996 or 2003). Any hearing aid that does not meet the specifications should be returned to the manufacturer for repair before being fitted to the user. Real-ear measures are recommended to ensure appropriate hearing aid gain. For compression hearing aids, the real-ear response is usually obtained at several levels (e.g., 50, 65, and 85 dB SPL) to make sure that measured gain is appropriate across a wide range of input levels. This procedure is important for both analogue and digital nonlinear hearing aids. The digital hearing aids are often programmed using the manufacturers' software, whose gain prescription has been reported to differ as much as 20 dB from measured insertion gain, especially at high frequencies (Hawkins and Cook, 2003; Keidser et al, 2003). Thus, real-ear measures are important to verify aided frequency responses.

Outcome measures are also recommended to validate hearing aid performance. Different outcome measures help to identify problems patients may be experiencing related to their hearing aids (Beck, 2000). Two clinical outcome measures that assess hearing aid benefit and user satisfaction are the Abbreviated Profile of Hearing Aid Benefit (APHAB; Cox and Alexander, 1995) and the Satisfaction with Amplification in Daily Life (SADL; Cox and Alexander, 1999), respectively. The APHAB contains four subscales, which are Ease of Communication (EC), Reverberation (RV), Background Noise (BN), and Aversiveness of sound (AV). A benefit score is computed by comparing unaided and aided APHAB scores. There are four aspects of hearing aid satisfaction measured using the SADL. These include Positive Effect (PE), Service and Cost (SC), Negative Feature (NF), and Personal Image (PI). A Global Score (GS) is based on a combination of all four subscale scores. These outcome measures help to identify problems that patients experience so that clinicians can better solve them.

In addition to the aforementioned measures, several fitting protocols recommend obtaining other unaided and/or aided loudness measures (ASHA, 1998; Kochkin, 2003; Mueller, 2003; Washington University School of Medicine, 2004). Generally, a measure of unaided loudness

is referred to as a “loudness discomfort level” (LDL). This measure identifies the loudest level a patient can tolerate without feeling uncomfortable. LDLs have been shown to be important for proper setting of the maximum output of hearing aids (Walker et al, 1984; Hawkins et al, 1987, 1992). Aided loudness perception is measured to ensure that a patient perceives soft, moderate, and loud sounds appropriately with compression hearing aids (ASHA, 1998). For example, a 65 dB SPL input should sound “comfortable,” and an 85 dB SPL input should sound “loud, but OK” to the patient.

Another test frequently recommended in a hearing aid fitting protocol is the aided speech recognition test in quiet and/or in noise (e.g., Humes, 1999; Taylor, 2003; Wilson, 2004). Some widely used speech-in-noise tests include the Connected Speech Test (CST; Cox et al, 1987), Hearing-in-Noise Test (HINT; Nilsson et al, 1994), and the Quick Speech-in-Noise (QuickSIN) Test (Killion et al, 2004). Of these tests, the QuickSIN is the fastest to administer. The QuickSIN can be used for three purposes (Taylor, 2003). First, it provides a measure of a patient’s “signal-to-noise ratio (SNR) loss,” when it is administered at a sufficiently high level (e.g., 75–80 dB HL). The loss in SNR provides an estimate of a patient’s unaided intelligibility in noisy situations. Patients with a moderate and severe SNR loss (e.g., >7 dB SNR) may require a very favorable SNR to understand speech in the presence of noise (Etymotic Research, 2001). These patients usually need additional counseling on how to cope with noisy listening situations. Measuring SNR loss also helps the patient set realistic hearing aid expectations. Second, QuickSIN can provide a measure of aided intelligibility in noise when speech is presented at a soft-to-moderate level (e.g., 45 dB HL). This measure allows both the clinician and patient to observe how much the hearing aid reduces the negative effects of noise. Small changes in intelligibility may require an adjustment of gain or further counseling. Finally, QuickSIN can be used to determine a patient’s candidacy for directional microphones and/or an FM system.

To include all of the above described

measures should increase the likelihood of a proper hearing aid fit and maximize hearing aid benefit. Other potential advantages may include providing evidence for third-party reimbursement, providing training opportunities for audiology students, and so on. However, clinicians rarely perform the basic recommendations (Mueller, 2003; Kirkwood, 2006). Medwetsky et al (1999) reported that fewer than half of hearing aid dispensers conducted verification procedures, and loudness measures and speech-in-noise testing were only performed by a limited number of clinicians. According to Mueller (2003), only 61% of audiology practices routinely conducted prefitting loudness measures (using speech instead of pure tone as the stimulus), and as high as 75% of the respondents in the survey never/seldom performed any speech-in-noise measures.

There may be many reasons for the low use of loudness measures and speech-in-noise testing, such as lack of equipment and training. One primary reason, however, is that these measures add time to the hearing aid fitting appointment. In addition, there is a lack of direct evidence that these measures will improve hearing aid benefit and/or user satisfaction. The effectiveness of loudness measures and speech-in-noise testing has not been consistently established (for review, see Killion and Gudmundsen, 2005; Mueller and Bentler, 2005). For example, no strong correlation was found between speech-in-noise tests and hearing aid user satisfaction (e.g., Humes et al, 2003; Walden and Walden, 2004). Patient reports of problems associated with loudness have been shown to be reduced when the measured real-ear saturation response (RESR) was below the patient’s measured LDL (e.g., Munro and Patel, 1998; Bratt et al, 2002). However, the extent to which LDL measures improve hearing aid benefit and/or user satisfaction needs further research (Mueller and Bentler, 2005).

In the current study, a randomized controlled single-blind design was used to compare two hearing aid fitting protocols. Both protocols consisted of basic measures recommended by all published guidelines, such as 2-cc coupler responses and real-

ear measures, but only one of the two protocols included loudness and speech-in-noise measures. The two protocols were compared to determine whether the inclusion of loudness and speech-in-noise measures would result in improved hearing aid benefit and user satisfaction, as assessed by the APHAB and SADL, respectively. We expected that the protocol with the additional measures would lead to higher benefit and satisfaction scores because the extra information would allow for a better hearing aid fitting. Specifically, the unaided and aided loudness measures could be used to determine the maximum output and/or adjust gain (e.g., Valente and Van Vliet, 1997). The result of the QuickSIN test could be used to adjust the frequency response of the hearing aid when the aided QuickSIN scores showed minimum improvement over the unaided scores. More importantly, the results can be used as a good tool for counseling patients (Wilson, 2004). For example, patients could be encouraged to set realistic expectations about how their hearing aids would help them in background noise. Such counseling could improve the amount of hearing aid benefit and level of user satisfaction, as could be reflected on the APHAB and SADL, respectively.

## METHODS

### Subjects

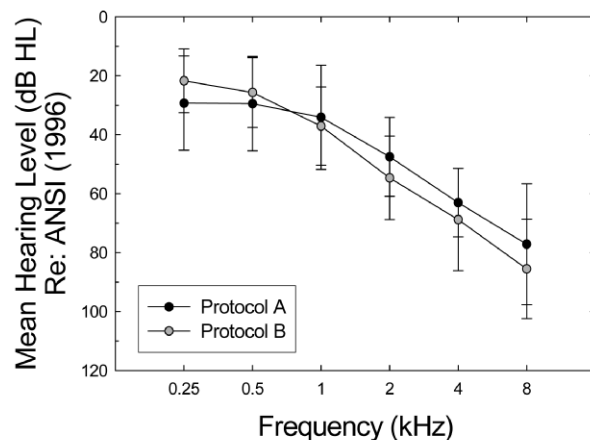
Participants included a total of 32 patients who came to the Gebbie Hearing Clinic at Syracuse University. Power analysis based on APHAB and SADL normative data, as published in Cox and Alexander (1995, 1999), indicated that the sample size was sufficient to achieve  $\beta = 0.80$  at  $\alpha = 0.05$ . All of the patients, who were recommended a midlevel digital behind-the-ear (BTE) or in-the-ear (ITE) hearing aid (see below for specific models) during the recruitment period for this study, were invited to participate in this study. All but one of the patients who met the criteria within this time frame agreed to participate (participation rate = 96.97%) in this study. Patients all signed a consent form approved by the

Institutional Review Board of Syracuse University prior to their participation. They made full payment for their hearing aids and received a 50% discount for related service charges.

Participants were informed that one of two methods would be used to fit their hearing aids. They had no knowledge of what these methods were or how they differed from each other. The first participant to accept the offer to participate in the study was randomly assigned to one of the two groups (which happened to be Group A). The following participants were alternately assigned to Group B and Group A when they consented to be in the study. Participants assigned to Group A were fitted with hearing aids using Protocol A, and participants in Group B were fitted using Protocol B. For both groups, participants' age and gender are shown in Table 1, and the degree of their hearing loss is shown in Figure 1. A Mann-Whitney *U*-test indicated that

**Table 1. Participants' Age and Gender**

	Group A	Group B
<b>Age (years)</b>		
Median	78.0	76.5
Mean	74.5	74.4
Standard Deviation	11.3	9.9
<b>Gender</b>		
Male	10	8
Female	6	8



**Figure 1.** Mean pure-tone thresholds re: ANSI (1996) for participants in Group A (black circles) and B (grey circles). Only thresholds for the ears that were to be aided are shown. The error bars represent one standard deviation.

there were no significant differences between the two groups of patients' ages ( $U = 271.500$ ,  $p = 0.792$ ) and degrees of hearing loss ( $p > 0.05$ ). However, the difference in thresholds between the two groups approached statistical significance at 2000 Hz ( $t = -1.962$ ,  $p = 0.055$ ), with participants in Group B having poorer thresholds than participants in Group A.

Eleven of the 16 participants in Group A and 10 of the 16 participants in Group B were fitted with hearing aids for the first time. The other participants in the two groups had comparable hearing aid experience (i.e., 5–20 years). Table 2 identifies the specific hearing aids the participants were fitted with. The make, model, and style of the hearing aids were not controlled for but were comparable between the two groups. Specifically, the midlevel digital hearing aids fitted were GN ReSound Canta2 and 4 series, Oticon Gaia, Phonak Aero, and Unitron Unison6. They all have three programs and similar standard features such as directional microphone, feedback cancellation, digital noise reduction, and telecoil.

## Procedures

Data were collected over a total of six visits to the clinic. These visits consisted of a prefitting consultation, hearing aid fitting, three postfitting follow-ups within the 45-day trial period, and one last visit scheduled at three months postfitting. Three certified clinicians at the Gebbie clinic (i.e., the first, third, and fourth authors) conducted the hearing aid fitting. The clinicians performed the tests according to the group the participants

were assigned. To minimize biases, the clinicians had no control in assigning participants and did not have access to the results from the outcome measures before the study was concluded.

## Prefitting Consultation

During a two-hour prefitting consultation, pure-tone and speech audiometric test results were reviewed with the participant; different types of hearing aid styles, models, and features available for the participant's degree of loss were discussed; and warranty and trial period information were provided. Also, an impression of the ear was taken, and the unaided section of the APHAB was completed. The clinician guided the participants through the first two items of the APHAB questionnaire and answered any questions raised by the participants. The clinician then left the room, and the participants completed the questionnaire.

For Group B participants, two more tests were conducted during the prefitting consultation, the QuickSIN and the LDL. The QuickSIN was administered at 70 dB HL for participants who had a speech recognition threshold (SRT) of <45 dB HL or at 40 dB SL re: SRT for participants who had an SRT of >45 dB HL. One QuickSIN list (six sentences) was presented to each of the participant's ears separately via a TDH-49 headphone (Telephonics, Farmingdale, NY). The sentences in QuickSIN are prerecorded at six SNRs, from 25 to 0 in steps of 5 dB SNR. Participants' scores were based on the number of key words in the sentences they correctly repeated. The final score represented the participant's total SNR loss. Participants with severe SNR loss ( $\text{SNR} > 7$ ) (Etymotic Research, 2001; Taylor, 2003) were counseled extensively on noise reduction technologies and rehabilitative strategies during the prefitting consultation and the hearing aid fitting. The QuickSIN provided the participant experience listening to speech in noise and the clinician a quantitative measure of the participant's ability to process speech in noise. These allowed for a more individualized counseling session on explaining strategies for listening in noise and setting realistic expectations for

**Table 2. Hearing Aids Used in the Study**

	Group A	Group B
<b>Make</b>		
Oticon	1	1
Phonak	1	1
GN ReSound	11	10
Unitron	3	4
<b>Style</b>		
BTE	12	12
ITE	4	4
<b>Binaural/Monaural</b>		
Binaural	12	13
Monaural	4	3

aided benefit in noisy situations. This was in contrast to the general instruction on the function and usage of directional microphones provided to Group A participants.

The second measure conducted for Group B was the LDL. LDL was obtained under headphones at 500 and 2000 Hz for each ear separately, using the method described by Hawkins et al (1987). Briefly, participants judged the loudness of a 500 and 2000 Hz narrowband noise using a scale that had nine categories, ranging from “very soft” to “painfully loud.” The clinician zeroed in at the “uncomfortably loud” level and, using a bracketing procedure, reached the level where “loud, but OK” was reported. This level was taken as the participant’s LDL for that specific frequency. Previous studies have recommended 500 and 3000 Hz (Bentler and Cooley, 2001), 500 and 4000 Hz (Hawkins et al, 1987), or 500, 1000, and 2000 Hz (Liu and Chen, 2000) to obtain the LDL. In the present study, 2000 Hz was chosen because all the manufacturers’ fitting software included a 2000 Hz band for gain manipulation, whereas some software (e.g., Phonak PFG v8.3, GN ReSound Aventa v1.50) did not have a 3000 Hz band. Including 500 and 2000 Hz made it easier to adjust the gain of the aid in response to the LDL. The LDL test ear, test frequency, and presentation level were randomized across participants. LDL information was utilized in the hearing aid fitting as described in the next section.

### Hearing Aid Fitting

Two weeks after the prefitting consultation, participants returned to the clinic for their hearing aid fitting. Prior to this visit, an electroacoustic analysis was conducted to verify that the hearing aids were performing within the manufacturers’ specifications (ANSI, 1996). All hearing aids met the specifications.

At the hearing aid fitting visit, the hearing aids were programmed using the NOAH system (HIMSA, St. Paul, Minnesota) through the Hi-Pro interface. Regardless of the make and model of the hearing aid, the omnidirectional mode was always programmed as the default for the first program; the directional-microphone mode was typically pro-

grammed in the second program; and the third program was reserved for telephone use. Gain was prescribed based on a participant’s recent audiogram and, for Group B participants, also their LDLs, using the Desired Sensation Level [input/output] (DSL[i/o]) fitting rationale (Cornelisse et al, 1995) in NOAH. If DSL[i/o] was not available in the manufacturer’s fitting software (mainly GN ReSound Aventa v1.50), then the manufacturer’s recommended fitting rationale was used.

After programming the aid, real-ear measures were obtained using the Verifit system (Audioscan, Dorchester, Ontario). The target responses were set for the DSL[i/o] rationale. Although the DSL[i/o] rationale is typically used for children, it is also used for adults (for review, see Scollie et al, 2005). The system was calibrated following the procedure specified on the Verifit user’s manual (Etymonic Design Incorporated, 2005). Briefly, the open end of the probe tip was placed over the reference microphone, and the probe module (i.e., the probe and the reference microphone) was held five inches in front of the Verifit speaker. A short tone sweep was presented to obtain the frequency response of the probe with the goal to calibrate it to the flat frequency response of the reference microphone.

The Speechmap module was activated to obtain the real-ear aided response (REAR). In this module, a discourse spoken by a male voice was presented at 55, 70, and 75 dB SPL, representative of soft, average, and loud speech level (Etymonic Design Incorporated, 2005). The participant was seated approximately three feet in front of the Verifit speaker. Hearing aid gain was adjusted via NOAH so that the output for the 70 dB SPL input matched the 70 dB SPL DSL[i/o] targets on the Speechmap. It was difficult to meet the targets for all frequencies, especially for participants with severe high-frequency hearing loss or significant loudness recruitment. In these cases, gain was adjusted in NOAH to ensure that at least the output for the 55 dB SPL input was above the participant’s thresholds. It should be noted that average real-ear unaided response (REUR) provided by Verifit was included in all real-ear measures

reported in this study (Etymonic Design Incorporated, 2005).

The LDL information was primarily used to set the saturation sound pressure level (SSPL-90) in the hearing aids (Hawkins et al, 1987, 1992). Also, the addition of LDL information changed the DSL[i/o] target across frequencies. Table 3 shows the difference in the mean DSL[i/o] target output with and without the measured LDLs at 500 and 2000 Hz. The difference was significant at all frequencies except 3000 and 6000 Hz.

### Postfitting Follow-Ups

Participants were scheduled to return to the clinic every two weeks after their initial hearing aid fitting. Thus, participants were scheduled for three visits within their 45-day trial period. During each visit, hearing aid adjustments were made based on participants' comments. For example, if a participant complained that the hearing aid sounded "too loud," then the gain was reduced. If the participant felt that he or she was listening "in a barrel," then the gain in the low-frequency region was adjusted.

At the first postfitting visit, Group B participants were given two additional tests: the QuickSIN and aided loudness tests. The QuickSIN was administered unaided and aided at 50 dB HL in the sound field. The results were primarily used for counseling purposes. For example, participants with poor aided QuickSIN scores were reinstructed on how to maximize their directional microphones (i.e., identifying noisy listening situations, switching between programs, etc.) and how to use other strategies for listening in noise (i.e., sitting close to the

talker, practicing speech reading, and asking the talker for paraphrasing and clarification, etc.). The low-frequency gain was reduced, and mid- to high-frequency gain was increased for a small number of participants who showed little to no improvement in aided QuickSIN scores.

The second additional test given only to Group B was aided loudness. This test was evaluated using a speech-shaped noise in the sound field. The noise was generated by a GSI-61 audiometer (Grason-Stadler, Madison, Wisconsin) and presented at 0° azimuth. Participants' loudness perception was obtained binaurally at 50, 65, and 85 dB SPL in the sound field, following the procedure recommended by Washington University School of Medicine (2004). Adjustments of gain were made to the hearing aids if the participant failed to report "soft, but audible," "comfortable," and "loud, but OK" in response to the 50, 65, and 85 dB SPL presentation level, respectively. To administer this test through the GSI-61 audiometer, these values were converted to 30, 45, and 65 dB HL using ANSI (1989). These presentation levels were randomized across participants. It should be noted that the sound pressure level measured at the location of the hearing aid microphone was 2–3 dB lower than the presentation levels of 50, 65, and 85 dB SPL. However, all the participants in Group B reported that the softest input was audible to them.

At the last of the three postfitting visits, Group A and B participants were asked to complete the aided section of the APHAB and SADL. The clinicians were not present while these questionnaires were being filled out. To estimate longer-term benefit and satisfaction, participants

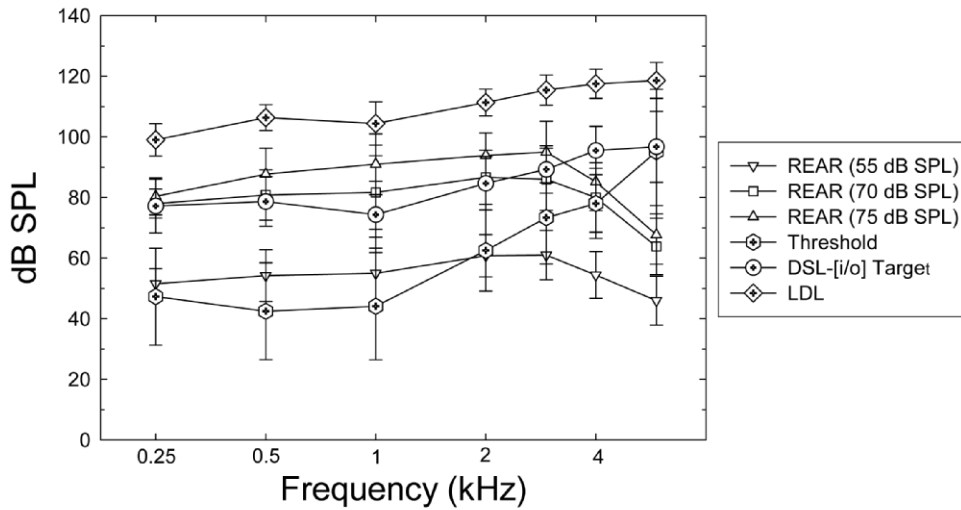
**Table 3. Differences across Frequencies in Mean DSL[i/o] Target Output as a Result of Incorporating Measured LDLs at 500 and 2000 Hz**

	250 Hz	500 Hz	1000 Hz	2000 Hz	3000 Hz	4000 Hz	6000 Hz
Mean Target Output (with Predicted LDL)	71.000	76.074	77.519	91.741	96.647	99.741	106.704
Mean Target Output (with Measured LDL)	73.185	77.481	79.741	93.222	97.471	101.222	107.593
Difference	2.185	1.407	2.222	1.481	0.824	1.481	0.889
<i>t</i> value	-7.49	-3.683	-5.547	-3.730	-1.171	-2.770	-1.986
<i>p</i> value	<0.001	0.001	<0.001	<0.001	0.259	0.010	0.058

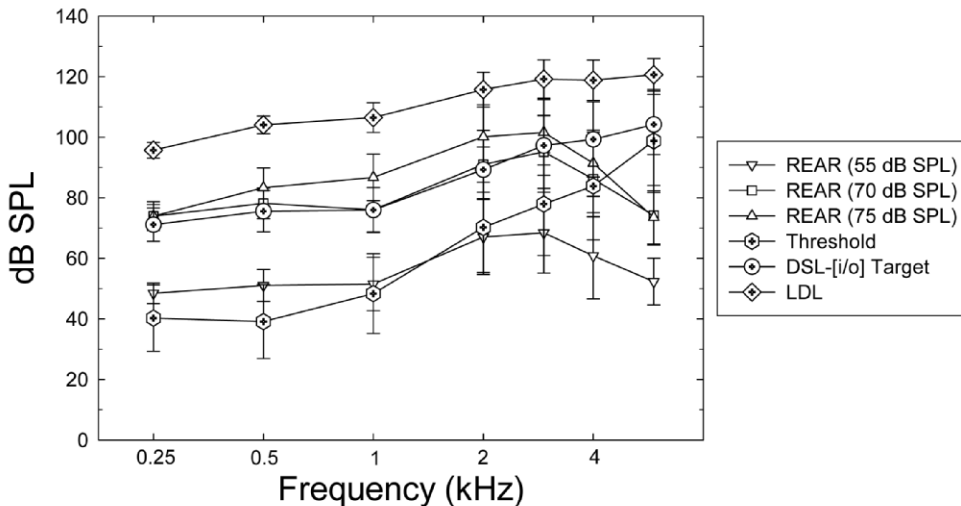
were asked to come back to the clinic at three months postfitting to complete the APHAB and SADL questionnaires again. The questionnaires were mailed to participants who could not return to the clinic. Mail surveys have been shown to not bias the hearing aid user as much as phone interviews (Dillon et al, 1991). A phone interview was conducted for only five participants who did not return their surveys by mail. Three graduate students in audiology conducted these phone interviews and were not given any information about the study.

**RESULTS**

REARs were obtained for a 55, 70, and 75 dB SPL input for both groups of participants (Figures 2 and 3). All participants' REARs for the 70 dB SPL (average) input met the DSL[i/o] targets up to 3000 Hz. Above 3000 Hz the REARs were lower than the targets. The REARs for the 55 dB SPL (soft) input were above all the participants' pure-tone thresholds up to 2000 Hz but below thresholds for participants whose hearing loss was more severe at the higher frequencies. The REARs for



**Figure 2.** Real-ear measure results for Group A participants. REARs were obtained for speech inputs at 55 (triangles down), 70 (squares), and 75 (triangles up) dB SPL. Also shown are the DSL[i/o] target (pluses in circles) and predicted LDL (pluses in diamonds). LDL was predicted by the Verifit system. The error bars represent one standard deviation.



**Figure 3.** Real-ear measure results for Group B participants. REARs were obtained for speech inputs at 55 (triangles down), 70 (squares), and 75 (triangles up) dB SPL. Also shown are the DSL[i/o] target (pluses in circles) and measured LDL (pluses in diamonds). LDL was obtained by entering LDL values at 500 and 2000 Hz into the Verifit system. The error bars represent one standard deviation.



the 75 dB SPL (loud) input were below the LDL values, ensuring that gain was not too high to cause loudness discomfort.

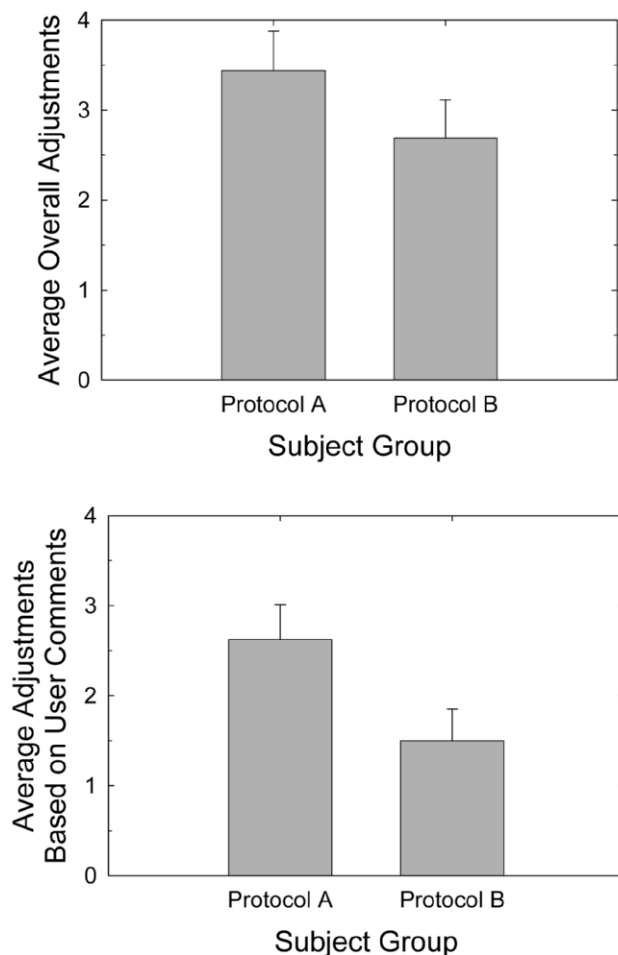
The total number of gain adjustments made within the 45-day trial period is shown in the top panel of Figure 4. These include adjustments made to the participants' aids based on their comments, real-ear measures, and, for Group B, also the QuickSIN and LDL information. The number of gain adjustments made based only on the participants' comments is shown in the bottom panel of Figure 4. Overall, more total adjustments were made for participants in Group A than in Group B, but the difference was not statistically significant ( $t = 1.229$ ,  $p > 0.05$ ). However, significantly more adjustments were made for Group A than Group B based only on the participants' comments ( $t = 2.149$ ,  $p = 0.040$ ). Interestingly, the

complaints made by several of the Group A participants were related to noise and loudness, whereas few such complaints were made by Group B participants. This finding suggests that participants in Group B had more realistic expectations about using their hearing aids in noise and/or their hearing aids were more fine-tuned than Group A because of the LDL and QuickSIN measures. As a result, Group A participants had requested more adjustments in the follow-up visits.

The difference between the initial (on the fitting day) and final (at the end of the 45-day trial period) hearing aid gain was determined by subtracting the users' initial gain values displayed on NOAH from their final gain values displayed on NOAH for a 50 and 80 dB SPL input. The range of differences across frequency was larger for participants in Group A (-10 to 26 dB) than participants in Group B (-10 to 15 dB). However, the differences were not statistically different for any frequencies ( $p > 0.05$ ).

For Group B participants, QuickSIN scores significantly improved in the aided condition compared to the unaided ( $t = 5.167$ ,  $p < 0.001$ ). The average QuickSIN unaided score was 14.75 dB SNR (3.5–25.5 dB SNR), and the aided score was 9.375 dB SNR (1.5–20.5 dB SNR). The LDLs across all frequencies measured for Group B participants were significantly higher (1.5–2 dB) than those estimated by the Verifit system ( $p < 0.003$ ).

Both groups of participants completed the APHAB and SADL questionnaires at the end of the 45-day trial period and again at three months postfitting. However, only 11 of the 16 participants in Group A and 14 of the 16 participants in Group B completed the two questionnaires at the three months postfitting. Of the five participants who dropped out in Group A, one switched to another model at the end of the trial period, two did not complete the follow-up for health reasons, and two were displeased with the performance of the hearing aids and did not wish to complete the three-month follow-up. Of the two participants who dropped out in Group B, one did not respond to the mail survey or phone interview, and the other upgraded his or her hearing aids,



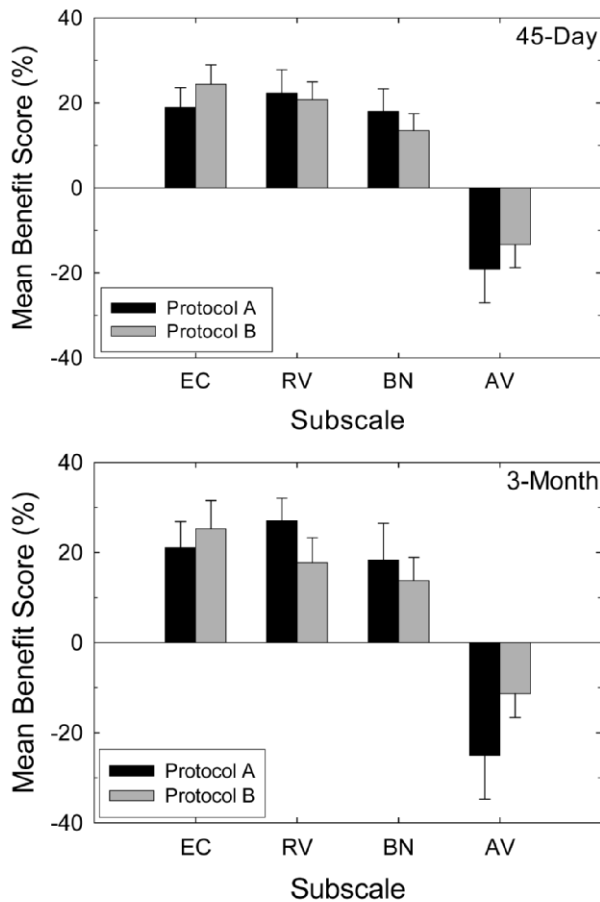
**Figure 4.** Number of hearing aid adjustments made within the 45-day trial period. Top panel: total number of adjustments; bottom panel: number of adjustments based on users' comments. The error bars represent one standard error.

which disqualified him or her from the study.

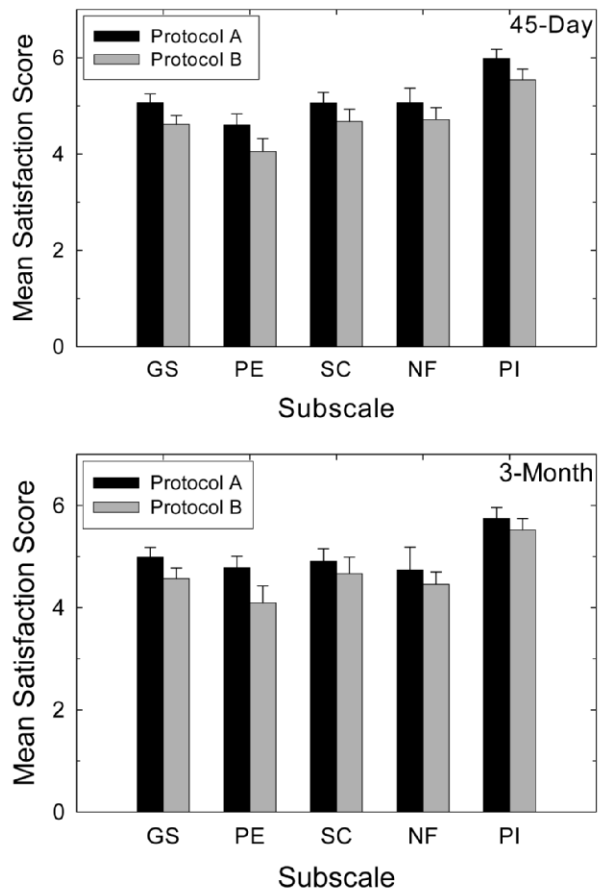
Results from the APHAB and SADL questionnaires, measured at 45 days and three months postfitting, are shown for both groups of participants in Figures 5 and 6, respectively. Note that participants who did not complete the follow-up three months postfitting were not included in the bottom panels of these figures. No significant differences ( $p > 0.05$ ) were observed between Group A and B participants' APHAB and SADL subscale scores at the end of the 45-day trial period or three months postfitting.

A paired  $t$ -test revealed that there were no significant differences between Group A participants' scores on the APHAB at 45 days and three months postfitting

( $p > 0.05$ ). However, Group A participants' scores on the SC and NF subscales of the SADL were significantly lower at three months compared to 45 days (SC:  $t = 2.701$ ,  $p = 0.022$ ; NF:  $t = 2.250$ ,  $p = 0.048$ ) (Figure 7). The Wilcoxon signed rank test was used to analyze the PI subscale scores because the distribution of the data failed the normality test. The PI score was significantly lower for Group A participants at the end of three months compared to 45 days ( $W = -34.00$ ,  $p = 0.016$ ). No significant within-group differences on the APHAB and SADL were observed for Group B participants at 45 days and three months. Thus, satisfaction remained constant for participants in Group B but decreased over time on some subscales for participants in Group A.



**Figure 5.** Mean APHAB benefit scores for Group A (black bars) and B (grey bars) participants at 45 days postfitting (top panel) and at three months postfitting (bottom panel). The data in the bottom panels were based on 11 and 14 participants in Group A and B, respectively. The error bars represent one standard error.



**Figure 6.** Mean SADL satisfaction scores for Group A (black bars) and B (grey bars) participants at 45 days postfitting (top panel) and at three months postfitting (bottom panel). The data in the bottom panels were based on 11 and 14 participants in Group A and B, respectively. The error bars represent one standard error.

## DISCUSSION

In the current study, two different hearing aid fitting protocols were evaluated in how they affect hearing aid benefit (APHAB) and satisfaction (SADL). Protocol A and B included 2-cc coupler responses and real-ear measures. The two protocols differed in that Protocol B included a speech-in-noise test (QuickSIN test) and loudness measures (LDL and aided loudness) and Protocol A did not. The participants who were fitted with the two different protocols had similar degrees of hearing loss, age, gender, and experience with hearing aids.

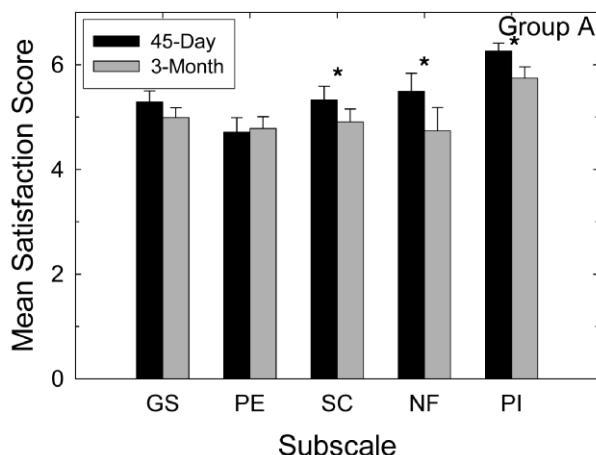
There were no significant differences in the participants' APHAB and SADL subscale scores fitted with the two different protocols when tested at 45 days and at three months postfitting. However, the SC, NF, and PI subscales of the SADL yielded significantly lower satisfaction scores for Group A participants at the end of three months compared to 45 days, whereas satisfaction remained constant across all the SADL subscales for participants in protocol B.

One possible reason that minimal differences were observed between the two protocols is that both protocols included real-ear measures. At least one adjustment was made to every hearing aid based on the real ear measure. Yet it is still difficult to know whether, if one of the protocols had not included real-ear measures, it would have resulted in more changes based on patient comments and/or a final frequency response less similar to that of the protocol that did include the real-ear measure. Clearly, a study assessing the specific effect of real-ear measure on fitting protocols would be interesting. A recent survey reported that fewer than 50% of audiology practitioners perform real-ear measures half of the time or more (Kirkwood, 2006). Direct evidence in support of the relationship between real-ear measure and hearing aid benefit and satisfaction may help increase the number of clinicians who are investing their time performing this test.

Findings similar to the APHAB scores reported in this study were also observed in the Cunningham et al (2001) study. They used several outcome measures to

evaluate aided benefit (i.e., APHAB) and satisfaction in two groups of hearing aid users. Both groups were fitted with the same hearing aids in the same manner, including real-ear measures. The difference was that the hearing aids for one group were fine-tuned over a five-month period based on patient comments. The amount of change in gain was comparable to the gain change we reported in this study. Interestingly, as in our study, Cunningham et al (2001) found no statistically significant differences on any of the benefit and satisfaction measures. Hence, small amounts of change in gain do not lead to measurable increases in perceived benefit and satisfaction in hearing aid users.

Although most of the basic demographics were similar across the two groups of participants, other characteristics, such as physical, psychological, and medical conditions and expectation of hearing aids and motivation to try hearing aids, which have all been shown to affect hearing aid benefit and user satisfaction (e.g., Kricos et al, 1991; Gatehouse, 1994; Wilson and Stephens, 2002; Arlinger, 2003; Humes et al, 2003; Cox et al, 2005), were not controlled for in this study. The two participants in Group A who dropped out of the study because they were displeased with their hearing aids would have likely obtained very unfavorable satisfaction scores at three months postfitting. As a result, Group A participants' average APHAB and SADL scores may have been



**Figure 7.** Mean SADL satisfaction scores obtained at 45 days (black bars) and three months (grey bars) post-hearing aid fitting for Group A participants. The error bars represent one standard error. The \* symbol indicates significant difference between two SADL scores ( $p < 0.05$ ).

worse if these participants had stayed in the study.

Participants in protocol B showed a 5–6 dB SNR improvement in their aided QuickSIN scores compared to their unaided scores, which is similar to the SNR changes (4–6 dB) reported in Figure 4 of the Walden and Walden (2004) study. Despite the consistency in the results with previous studies, how the QuickSIN was used in this study might have contributed to small differences in benefit and satisfaction observed between the two fitting protocols. The number of hearing aid adjustments made based on QuickSIN was negligible because there are no published recommendations for how gain should be adjusted based on a QuickSIN score. Our adjustment and reprogramming of the hearing aids based on QuickSIN may not produce enough of a change in gain or frequency response to affect hearing aid benefit. We did, however, consistently use the QuickSIN scores to counsel participants, which may account for the stable satisfaction scores observed in Group B participants and not in Group A participants. Also, the QuickSIN test was initially intended in the prefitting to better determine how strongly to recommend directional microphones. However, due to the wide availability and affordability of directional microphones in contemporary digital hearing aids (e.g., entry-level models such as GN ReSound Canta2 and Unitron Unison3), participants in both groups were fitted with directional microphones. Recently, inconsistencies were reported across some QuickSIN lists by McArdle and Wilson (2006). This issue may also have affected the results in the present study because, if a difficult list was used in the unaided condition and an easy list was used in the aided condition, then it may have exaggerated the effectiveness of the hearing aids.

Despite these potential confounding factors, the participants fitted with Protocol B were more consistently satisfied with their hearing aids over a three-month period than participants fitted with Protocol A. Perhaps the QuickSIN and LDL information, obtained in Protocol B, gave clinicians an opportunity to provide these participants with more

focused counseling in areas such as background noise, feedback issues, and loudness of amplified sound. Indeed, Group B participants showed stable satisfaction scores on the SADL SC subscale, which measures satisfaction with the clinician's competency in service, and the NF subscale, which measures satisfaction with the clinician's ability to solve problems and respond to complaints, but scores on these two subscales for the participants in Group A decreased over time.

The number of hearing aid adjustments that were made to the hearing aids based only on the participants' comments was significantly less for Protocol B than Protocol A. This result was likely attributed to the LDL measure. When the participants' LDL values were used, it significantly changed the DSL<sub>[i/o]</sub> targets across frequencies. Similar findings have been reported in other studies (e.g., Bentler and Cooley, 2001). It may be argued that the small change in gain that we observed is not clinically important, but if it reduced the number of adjustments made to the hearing aids, it could be meaningful. Exclusive reliance on a patient's comments may result in over- or underamplification and, more importantly, if the patient needs to keep describing how unsatisfied he or she is with the quality of the amplified sound, he or she may lose confidence in the fitting procedure and/or the clinician. Even if after multiple adjustments the patient becomes satisfied with the hearing aid performance, there may be residual effects on his or her overall satisfaction. It should be noted, however, that the significant difference in adjustment between the two protocols has a low statistical power ( $\beta = 0.44$ ) and, therefore, more data would be needed to test this hypothesis.

In summary, inclusion of LDL, aided loudness, and speech-in-noise measures in a hearing aid fitting protocol tended to reduce the total number of hearing aid adjustments and significantly reduced the number of adjustments based on patient comments. Although these measures did not improve initial hearing aid benefit and satisfaction, patients who did not receive them showed a significant decrease in their hearing aid satisfaction over time.

**Acknowledgments.** The authors would like to thank Lea Georgantas, Allison Singleton, Margaret Senior, and Amy Diehl for their help collecting the data.

## REFERENCES

- American National Standards Institute. (1989) *American National Standards Specifications for Audiometers* (ANSI-S3.6-1989). New York: American National Standards Institute.
- American National Standards Institute. (1996) *Specification for Hearing Aid Characteristics* (ANSI-S3.22-1996). New York: American National Standards Institute.
- American National Standards Institute. (2003) *Specification for Hearing Aid Characteristics* (ANSI-S3.22-2003). New York: American National Standards Institute.
- American Speech-Language-Hearing Association. (1998) Guidelines for hearing aid fitting for adults. *Am J Audiol* 7:5–13.
- Arlinger S. (2003) Negative consequences of uncorrected hearing loss—a review. *Int J Audiol* 42:2S17–20.
- Beck LB. (2000) The role of outcomes data to health-care resource allocation. *Ear Hear* 21:S89–96.
- Bentler R, Cooley L. (2001) An examination of several characteristics that affect the prediction of OSPL90. *Ear Hear* 22:3–20.
- Bratt GW, Rosenfeld MAL, Poek BF, Kang J, Williams DW, Larson V. (2002) Coupler and real-ear measurement of hearing aid gain and output in the NIDCD/VA hearing aid clinical trial. *Ear Hear* 23:308–324.
- Cornelisse LE, Seewald RC, Jamieson DG. (1995) The input/output formula: a theoretical approach to the fitting of personal amplification devices. *J Acoust Soc Am* 97:1854–1864.
- Cox RM, Alexander GC. (1995) The Abbreviated Profile of Hearing Aid Benefit (APHAB). *Ear Hear* 16:176–186.
- Cox RM, Alexander GC. (1999) Measuring Satisfaction with Amplification in Daily Life: the SADL scale. *Ear Hear* 20:306–320.
- Cox RM, Alexander GC, Gilmore CA. (1987) Development of the connected speech test (CST). *Ear Hear* 8:119S–126S.
- Cox RM, Alexander GC, Gray GA. (2005) Who wants a hearing aid? Personality profiles of hearing aid seekers. *Ear Hear* 26:15–26.
- Cunningham DR, Williams KJ, Goldsmith LJ. (2001) Effects of providing and withholding postfitting fine-tuning adjustments on outcome measures in novice hearing aid users: a pilot study. *Am J Audiol* 10:51–65.
- Dillon H, Koritschoner E, Battaglia J, Lovegrove R, Ginis J, Mavrias G, Carnie L, Roy P, Forsythe L, Towers E, Goulias H, Macaskill F. (1991) Rehabilitation effectiveness II: Assessing the outcomes for clients of a national hearing rehabilitation program. *Aust J Audiol* 13:68–82.
- Etymonic Design Incorporated. (2005) *Audioscan Verifit® VF-1 Real-Ear Hearing Aid Analyzer User's Guide* (Version 2.5). Dorchester, ON: Etymonic Design Incorporated.
- Etymonic Research. (2001) *QuickSIN Speech-in-Noise Test* (Version 1.3). Elk Grove Village, IL: Etymonic Research.
- Gatehouse S. (1994) Components and determinants of hearing aid benefit. *Ear Hear* 15:30–49.
- Hawkins DB, Ball TL, Beasley HE, Cooper WA. (1992) Comparison of SSPL90 selection procedures. *J Am Acad Audiol* 3:46–50.
- Hawkins DB, Cook JA. (2003) Hearing aid software predictive gain values: how accurate are they? *Hear J* 56(7):26–34.
- Hawkins DB, Walden BE, Montgomery AA, Prosek RA. (1987) Description and validation of an LDL procedure designed to select SSPL90. *Ear Hear* 8:162–169.
- Humes LE. (1999) Dimensions of hearing aid outcome. *J Am Acad Audiol* 10:26–39.
- Humes LE, Wilson DL, Humes AC. (2003) Examination of differences between successful and unsuccessful elderly hearing aid candidates matched for age, hearing loss and gender. *Int J Audiol* 42:432–441.
- Keidser G, Brew C, Peck A. (2003) Proprietary fitting algorithms compared with one another and with generic formulas. *Hear J* 56(3):28, 32–38.
- Killion MC, Gudmundsen GI. (2005) Fitting hearing aids using clinical prefitting speech measures: an evidence-based review. *J Am Acad Audiol* 16:439–447.
- Killion MC, Niquette PA, Gudmundsen GI, Revitt LJ, Banerjee S. (2004) Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am* 116:2395–2405.
- Kirkwood DH. (2006) Survey: dispensers fitted more hearing aids in 2005 at higher prices. *Hear J* 59(4):40–50.
- Kochkin S. (2003) MarkeTrak VI: on the issue of value: hearing aid benefit, price, satisfaction, and repurchase rates. *Hear Rev* 10(2):1–8.
- Kricos P, Lessner SA, Sandridge SA. (1991) Expectations of older adults regarding the use of hearing aids. *J Am Acad Audiol* 2:129–133.
- Liu TC, Chen YS. (2000) Loudness discomfort levels in patients with conductive and mixed hearing loss. *Auris Nasus Larynx* 27:101–104.

McArdle RA, Wilson RH. (2006) Homogeneity of the 18 QuickSIN™ lists. *J Am Acad Audiol* 17:157–167.

Medwetsky L, Sanderson D, Young D. (1999) A national survey of audiology clinical practices. Part 1. *Hear Rev* 6(11):24–32.

Mueller GH. (2003) Fitting test protocols are “more honored in the breach than the observance.” *Hear J* 56(10):19–26.

Mueller GH, Bentler RA. (2005) Fitting hearing aids using clinical measures of loudness discomfort levels: an evidence-based review of effectiveness. *J Am Acad Audiol* 16:461–472.

Munro K, Patel R. (1998) Are clinical measurements of uncomfortable loudness levels a valid indicator of real-world auditory discomfort? *Br J Audiol* 32:287–293.

Nilsson M, Soli SD, Sullivan JA. (1994) Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and noise. *J Acoust Soc Am* 95:1085–1099.

Scollie S, Seewald R, Cornelisse L, Moodie S, Bagatto M, Lurnagaray D, Beaulac S, Pumford J. (2005) The Desired Sensation Level multistage input/output algorithm. *Trends Amplif* 9:159–197.

Taylor B. (2003) Speech-in-noise tests: how and why to include them in your basic test battery. *Hear J* 56(1):40, 42–46.

Valente M, Van Vliet D. (1997) The independent hearing aid fitting forum (IHAF) protocol. *Trends Amplif* 2:6–35.