Introduction

In the current era of audiologic management, much of the focus has been placed on hearing aid technologies and fitting schemes. With rapid advances in component miniaturization and digital processing algorithms, there is a tacit assumption that all hearing loss can be managed. Yet, any clinical audiologist will attest to the fact that success for one individual is no guarantee of success for another, even when audiometrically similar. Even the earliest of audiological efforts suggested there was little relationship between the audiogram and speech perception ability, pre- or post-fitting (Hirsch 1952). The earliest “fitters” viewed the audiometric profile as useful only as a guide to how powerful a hearing aid should be, not as a guide to the best frequency characteristics (Davis 1947). With the current programmability across many frequency bands, and the multiple approaches to assessing outcomes of amplification fittings, one would expect the relationship between the audiometric data and the hearing aid effectiveness data to be more discernable.

How Do Thresholds Relate to Success?

There has been little updating over the past 70 years of the threshold-taking procedures that go on in audiology clinics of the world. The early investigators/clinicians eventually concurred on what constitutes normal audiometric thresholds (e.g., refer to Stevens 1951). The first ASA standard providing specifications for audiometers, masking noise, and even shock hazard, appeared in 1951 (ASA 1951), some 30 years after the first prototype of a modern, vacuum-tube audiometer was developed. The Hughson-Westlake approach to establishing threshold values, first published in 1944, continues to be the psychophysical approach of choice since Carhart and Jerger (1959) argued that all audiologists should follow the same methodology so that test results would not be influenced by differences in test procedures.

Outcome measures, on the other hand, continue to multiply. Aided speech perception tests and self-report inventories are the tools most often used to evaluate the success of hearing aid fittings. Self-report outcome measures used to quantify the success from hearing aid fittings cover many domains, including handicap (activity limitation), disability (participation restriction), device use, satisfaction, and quality of life, along with economic considerations such as cost-benefit and cost-utility. For the purposes of this paper, those studies relating audiometric correlates to hearing aid success in any of the domains will be considered. One should consider, however, that the bandwidth, distortion and overall fidelity of earlier technology may have precluded meaningful outcomes (Bentler and Duve 2000).

In 1982, Tom Giolas did a “critical review of the evidence” (in evidence based practice [EBP] terminology) to explore the relationship between audiometric variables (pure tone thresholds and speech discrimination scores) and self-reported hearing handicap. His goal and that of other contemporary clinician-researchers of the day was to attempt to explain the variance in hearing handicap using audiometric measures. The correlations between pure tone thresholds and self-report data ranged from .37 to .73; the correlations between speech recognition and self-reported handicap ranged from 89
-12 to -73. Attributing these weak, but widely variable, correlations to the various inventories used and the varied subject populations studied, such effort only reinforced Davis (1947) and Hirsch’s (1952) early contentions of the poor relationship between objective and subjective measures of hearing and hearing aid success that would be further interrogated over the following 50-60 years. Although threshold measurement had reached some stage of standardization by that time, outcome measures using self-report were only beginning to provide psychometric data to support the reliability and validity of their usage. Several of those inventories are still in use today and will be discussed here.

In 1983, two studies were published addressing the relationship between thresholds and hearing aid success. Weinstein and Ventry (1983) studied the correlation of pure tone threshold average (PTA) to scores on the Hearing Handicap Inventory for the Elderly (HHIE), a commonly used outcome measure today, and noted a .61 correlation over 100 non-institutionalized older adults (mean age 75.7). Two years later Brainerd and Frankel (1985) reported a correlation of .35 between PTA and the Social Hearing Handicap Index (SHHI), probable cause of that inventory’s demise from our repertoire. The SHHI was one of two inventories studied by these investigators; neither provided useful predictive power leading these investigators to assert, “…the relationships between audiometric (data) and self-report measures are weak and appear unlikely to be assessing the same issue (p. 92).” Similarly, only a moderate correlation (.54) between PTA and the Revised Hearing Performance Inventory was reported by Hawes and Niswander (1985). By 1992, Mulrow, Tuley and Agular reported that less than 11% of the variance in hearing aid success (defined as improvement in HHIE scores) could be accounted for by functional gain in the high frequencies and gain in Speech Reception Thresholds (SRTs), both directly related to threshold values. Of interest in that study of 39 adults (mean age of 52) with noise induced hearing loss, the authors concluded that baseline handicap (as measured with the HHIE), in combination with age, education, and number of medications, was more predictive of hearing aid success with or benefit from hearing aids than routine audiometric data.

Perhaps the best correlations of the era came from the work of Coren and Hakistan (1992). The investigators developed a Hearing Screening Inventory that produced a high correlation to PTA (.82) but was never implemented as an outcome measure for hearing aid success or benefit, in part, perhaps, because it was developed for group survey administration and not for outcomes use.

Satisfaction with hearing aids can certainly be construed as success. In their exhaustive review of studies relating degree of hearing loss to satisfaction measures, Wong, Hickson and McPherson (2003) noted that 9 out of 14 related studies found no such relationship. They suggest that such findings are related, in part, to the homogeneity of the subjects within a given study. Studies reporting on subjects with a wider range of hearing loss (e.g., Dillon, Biryles and Lovegrove 1999; Kochkin 2000; Gatehouse, Naylor and Eberling 2006) have reported moderate correlations between degree of loss and reported satisfaction.

In another large-scale study relating pre-fitting measures to satisfaction with amplification, Hosford-Dunn and Halpern (2001) correlated patient-related information (including gender, age, years of use, and pure-tone average) to global satisfaction measures obtained using the Satisfaction with Amplification in Daily Life (SADL; Cox and Alexander 1999). None of those variables alone or in combination with each other provided predictive insight; the highest significant correlation coefficient was for age (r = -1.18), indicating older subjects tended to report less overall satisfaction than younger subjects. Most recently, an attempt was made by Ricketts and Mueller (2000) to use high frequency average threshold and audiometric slope of hearing to predict success with directional microphone schemes implemented in modern hearing aids. Directional benefit was defined as the difference score based on Hearing in Noise Test (HINT; Nilsson, Soli and Sullivan 1994) scores obtained in omni- and directional modes. High frequency average loss was calculated as the average of thresholds at 2000, 3000 and 4000 Hz. No significant relationship emerged. Audiometric slope was calculated as the difference between thresholds measured at 2000, 3000, and 4000 Hz divided by 2.5 octaves (or the separation between 500 and 3000 Hz). Although the slopes across their 80 subjects varied considerably (from 18 dB/octave to -2.7 dB/octave), no significant correlation to directional benefit was found.

Perhaps the strongest evidence to support the relationship of hearing levels to hearing aid success can be inferred from Purdy and Jerram (1998). If hearing aid

---

1 Social Hearing Handicap Scale, from Ewertsen and Birk-Nielsen 1973.
2 Revised Hearing Performance Inventory, from Lamb, Owens and Shubert 1983.
use time is an acceptable outcome domain, these investigators found a significant correlation within their 102 subjects between hours of use and degree of hearing loss ($r = .42, p < .001$) and between hours of use and overall satisfaction ($r = .57, p < .001$). The authors conclude that subjects who were more satisfied tended to be those with greater hearing loss and, as a result, wore their hearing aids for more hours per day. These results are in contrast to other investigations involving use time and self-reported success (e.g., Crowley and Nabalek 1996; Cox 1997).

Recently, candidature for different signal processing schemes has been studied and audiometric correlates of preference and performance determined. Gatehouse et al. (2006) found that greater listening comfort as well as reported and measured speech intelligibility with linear hearing aids was associated with a flatter audiogram and wider dynamic range; higher listening comfort, higher satisfaction and reported intelligibility with nonlinear processors was associated with a more sloping audiogram and a more restricted dynamic range.

**What about Speech Measures?**

In 1952, Ira Hirsch noted, “So far as the relationship between the audiogram and discrimination loss is concerned, …here lies one of the greatest points of ignorance in contemporary audiology” (p. 149). If the pure tone information does not predict hearing aid success (and bears no relationship to the speech measures), maybe pre-fitting speech measures will offer some insight. The data suggest otherwise.

As was the case for PTA and hearing aid success reported earlier, Weinstein and Ventry (1983) also reported on speech perception ability as it related to handicap reduction (HHIE). They found a weak correlation ($r = -.42$) for their elderly subjects. Contrarily, Hawes and Niswander (1985) reported a higher correlation of word recognition to the Revised Hearing Performance Inventory (R-HPI) ($r = -.67$) than they had reported for pure tones. Bentler, Anderson, Niebuhr and Getta (1993a, b) found weak correlations between various measures of speech perception and self-report measures of hearing aid satisfaction and disability, among others, for a group of 65 aided adults followed over a one-year interval. In fact, no significant correlation could be found between traditional unaided speech measures and aided satisfaction in a critical review of the evidence undertaken by Killion and Gudmundsen (2005).

Occasionally, speech testing is used as a pre-test as well as an outcome measure. It should be recognized that almost any relationship of unaided word recognition ability to aided word recognition ability can be explained based on audibility. That is, if the testing is done before and after a careful fitting of amplification, one would assume the increased audibility would improve the measured performance. If the testing is accomplished in a difficult noise background (e.g., signal-to-noise ratios [SNRs] poorer than +5) that relationship may no longer hold, due to the reduced ability of the impaired auditory system to sort the speech from the noise. In fact, speech perception testing that is undertaken to determine the SNR for 50% performance may offer the greatest insight into the auditory integrity of the individual. Tests such as the Hearing in Noise Test (HINT; Nilsson et al. 1994) or the Speech in Noise (SIN) test (Fikret-Pasa 1993; Bentler 2000) are not based strictly on the audibility of the intended signal, but rather the integrity of the auditory filters to separate the intended signal from the background noise. At least one investigation has shown this test paradigm to provide predictive insight. Walden and Walden (2004) suggested that the QuikSIN (2001), a modified version of the SIN, was the best predictor of hearing aid success in daily life. In that study, a large number of predictive measures were included in a regression analysis, including age, PTA, audibility index (both aided and unaided), recorded NU-6 scores, QuickSIN (both aided and unaided) hearing aid experience, and current use time. Although the unaided audibility index bore a positive correlation to the aided audibility index ($r = .80$ at $p < .01$), none of the other audiometric measures held the same strong relationship. As for correlations between the predictive measures and the two outcome measures (IOI-HA3 and HAUS4), only hours of use ($r = .47$ to IOI-HA) and QuickSIN scores ($r = -.34$ to -.40 for both IOI-HA and HAUS) provided significant results.

**Other Audiometric Considerations**

When considering audiometric correlates to hearing aid “success”, the matter of dead regions needs to be addressed. Brian Moore describes these “regions of the cochlea with no (or very few) functioning inner hair cells
and/or neurons” as an old phenomenon (p. 46; Troland 1929), with significant consequence relative to hearing aid fitting (Moore 2001, 2004). Assuming we can actually identify those regions with considerable accuracy (a feat of controversy), it seems obvious that providing gain to areas void of transmission cells will not likely provide success to the hearing aid user. In fact, Preminger, Carpenter and Ziegler (2005) found that approximately one-third of those subjects with thresholds between 50 and 80 dB HL (re: ANSI 1996) tested positive for dead regions, using the Threshold Equalizing Noise (TEN) test (Moore 2001). Those subjects exhibited poorer sentence understanding in noise, and poorer self-perceived benefit while listening in backgrounds of noise or reverberation compared to those subjects who did not test positive for the condition. Using the same TEN test, data from Vestergaard (2003) support the notion that providing audibility within the dead region does not lead to concomitant increases in speech intelligibility.

It has been suggested that measures of loudness (comfort or discomfort) may also improve hearing aid success. In an evidence-based review of the literature, Mueller and Bentler (2005) asked the question: Are the clinical measures of LDL for adults predictive of aided acceptance and satisfaction of loudness for high inputs in the real world? Although over 200 articles addressed the topic, only three met the pre-defined criteria for inclusion. As the authors note, the evidence is weak due to the limited number of studies and the poor statistical power in the study, but does support the use of loudness comfort to increase user acceptance and success. Earlier work by Ricketts (1996) did not indicate that amplification outcomes (e.g., speech perception) would be enhanced by laboratory measures of loudness growth or preference (another type of audiometric testing). Similar results have been reported by Kliessling (2001) and Smeds et al. (2006).

Relative to predicting bilateral success with amplification, Mueller, Grimes and Jerome (1981) reported that the best predictor of binaural superiority over unilateral fitting of hearing aids was the slope of the performance-intensity function. In that investigation, both phonetically-balanced (PB) word lists and Synthetic Sentence Test (SSI) lists were presented to 24 older adults (mean age 69.4). The unilateral versus bilateral performance differences were unremarkable; yet, the slope of the PI function for PB words was a significant predictor of bilateral performance. With the current concern over binaural interference in the older adult, it seems logical that speech materials be tested in both unilateral and bilateral modes in order to better counsel those individuals for whom the unilaterally fit hearing aid will provide the better choice for communication in backgrounds of noise (Walden and Walden 2004; Sobiesiak and Bentler under review).

Understanding Those Factors Useful in Predicting Success

Some of the most comprehensive undertakings relative to understanding the relationship between audiological variables and hearing aid outcomes are attributable to Humes and his colleagues (Humes 1999, 2001; Humes, Wilson, Barlow and Garner 2002a; Humes, Wilson, Barlow, Garner and Amos 2002b; Humes, Wilson and Humes 2003). In the absence of any published large-scale studies aimed at modeling hearing aid “success” and identifying those factors that influence this success (e.g., Gatehouse, see Chapter 8 in this Proceeding), Humes (2003) has undertaken several factor-analyses efforts in an attempt to determine the number and nature of the dimensions of successful hearing aid outcome. Using a principal-components factor analysis, he examined the results of three studies of hearing aid outcomes: the NIDCD/VA multi-site study of 338 elderly users of single-channel, linear, and in-the-ear hearing aids; an Indiana University-based study of 53 elderly users of two-channel, wide dynamic range in-the-canal hearing aids; and, an Indiana University-based study of 173 elderly wearers of single-channel, linear, output compression-in-the-ear hearing aids. Factor analysis revealed that three to five principle components capture the individual differences in outcome measure. For the Indiana-based studies (n = 53 and n = 173) the three factors included speech discrimination scores, subjective benefit and satisfaction and hearing aid usage, and accounted for two-thirds of the total variance in the outcome measures. For the VA/NIDCD study, the five orthogonal components, or factors, included speech perception in quiet, speech perception in noise, subjective benefit, satisfaction, and hearing aid usage. For this analysis, aided speech perception in quiet accounted for 40% of the variance found in outcome measures, subjective benefit and satisfaction accounted for 22.8%, hearing aid usage accounted for 15.1%, aided speech perception in noise accounted for 6.7% and satisfaction accounted for 6% of the variance. The authors summarize their findings as:
1) Aided speech perception was best predicted by degree of hearing loss, cognitive performance and age; 2) Hearing aid use was best predicted by previous hearing aid use; 3) Hearing aid satisfaction was best predicted by aided sound quality measures.

In a recent and related study, Humes and his colleagues (Humes et al. 2003) sought to identify factors influencing the decision-making relative to amplification. Using three groups matched for age, degree/configuration of hearing loss, and gender proportion, all had significant hearing loss and were advised by an audiologist to pursue the use of hearing aids. One group (n = 26) opted not to pursue amplification (non-adherent group); one group (n = 24) opted to purchase the hearing aids but soon rejected them (reject-HA group); and, one group (n = 26) pursued hearing aids and were still using them at six months (adherent group). The group rejecting the hearing aid recommendation was not different from the other two groups in terms of any audiometric variable. For the two groups who tried the hearing aids, the only audiometric-related variable that distinguished the subjects was the higher pre-fitting loudness discomfort level (LDL) measures; finger dexterity was the other distinguishing predictor of hearing aid success. One might postulate that those with the better dexterity could more easily/quickly manage the volume control on the linear hearing aids worn by all subjects (adherents and hearing aid-reject groups).

Summary

It appears that the more traditional audiometric information such as thresholds, including degree and configuration of hearing loss, speech perception, and loudness ratings are of limited use in predicting hearing aid success, even when broadly defining success. It is clear that the greater the hearing loss, the more the use time of the amplification scheme. It is also clear from this review that other variables or factors account for the positive self-reported outcomes.

References


Dillon, H., Birtles, G., and Lovegrove, R. 1999. Measuring the outcomes of a national rehabilitation program: Normative data for the Client Oriented Scale of Improvement (COSI) and the Hearing Aid User's


Smeds, K., Keidser, G., Zakis, J., Dillon, H., Leijon, A.,
Grant, F., Convery, E., and Brew, C. 2006. Preferred
overall loudness: Listening through hearing aids in
field and in laboratory tests. *International Journal of

Sobiesiak, B., and Bentler, R. under review. Investigation
of binaural interference across two age groups.

New York: John Wiley and Sons.

Troland, L.T. 1929. The psychophysiology of auditory quali-
ties and attributes. *Journal of General Psychology*
2:28–58.

Vestergaard, M. 2003. Dead regions in the cochlea: Im-
plications for speech recognition and applicability of
the articulation index theory. *International Journal of

Walden, T.C. and Walden, B.E. 2004. Predicting success
with hearing aids in everyday living. *Journal of the

Weinstein, B.E., and Ventry, I.M. 1983. Audiometric cor-
relates of the Hearing Handicap Inventory for the

Hearing aid satisfaction: What does research from
the past 20 years say? *Trends in Amplification*