J Am Acad Audiol 8: 27–43 (1997)

Client Oriented Scale of Improvement (COSI) and Its Relationship to Several Other Measures of Benefit and Satisfaction Provided by Hearing Aids

Harvey Dillon*
Alison James†
Jenny Ginis‡

Abstract

Several methods for measuring the self-reported benefit and satisfaction provided by a hearing aid were compared by administering all methods to each of 98 subjects. Significant correlations between many of the measures and reasonably high test–retest correlations for two of the measures administered twice suggest that most of the measures provide valid estimates of benefit and/or satisfaction. One of the methods included is a new tool called the Client Oriented Scale of Improvement (COSI™). In this method, the client effectively writes the self-report questionnaire by nominating up to five listening situations in which help with hearing is required. At the conclusion of rehabilitation, reduction in disability and the resulting ability to communicate in these specific situations is quantified. Based on correlation analysis, the COSI method is as statistically valid as the much longer, more traditional questionnaires. Other features, such as relevance, diagnostic utility, compatibility with normal interviewing techniques, and good test–retest reliability, make it particularly suitable for routine clinical use.

Key Words: Benefit, handicap, hearing aids, satisfaction, usage

Abbreviations: 3FAHL = three frequency average hearing level; ASS = Aid Satisfaction Scale; COSI = Client Oriented Scale of Improvement; GAS = Goal Attainment Scaling; HAPI = Hearing Aid Performance Inventory; HAUQ = Hearing Aid Users' Questionnaire; HHIE = Hearing Handicap Inventory for the Elderly; MPHAB = Modified Profile of Hearing Aid Benefit; NAL = National Acoustic Laboratories (of Australia); PHAB = Profile of Hearing Aid Benefit; SHAPIE = Shortened Hearing Aid Performance Inventory for the Elderly; SNR = signal-to-noise ratio; SSS = Service Satisfaction Scale

The benefit provided by hearing aids and other aspects of a rehabilitation program can be assessed by one of two general methods. In the first method, benefit can be defined as the increase in speech understanding that wearing the hearing aid gives to the client. When defined in this way, benefit can be objectively measured using speech tests. In the second general method, clients can be questioned about the benefit they perceive in their everyday lives. This paper is about the second of these methods, but we will first briefly discuss the advantages and disadvantages of using objective speech tests.

The major advantage of speech tests for assessing benefit is that they objectively measure, with a known reliability, an increase in speech understanding, and it is the inability to adequately understand speech that most motivates people to obtain a hearing aid. The major disadvantage of speech tests for assessing aid benefit is that the results obtained depend strongly on the measurement conditions employed. Large improvements in scores can be obtained if the tests are administered in

*National Acoustic Laboratories, Chatswood, Australia; †Australian Hearing Services, Queensland, Australia; ‡Australian Hearing Services, Victoria, Australia

COSI™ is a trademark of Australian Hearing Services.

Reprint requests: Harvey Dillon, National Acoustic Laboratories, 126 Greville St., Chatswood, 2067, Australia

quiet, at low presentation levels, because the hearing aid can greatly increase the audibility of such speech. These benefits decrease and may disappear or even become negative if the same test material is presented in noise at high presentation levels, because audibility may then be mostly determined by the signal-to-noise ratio (SNR) at each frequency. Because clients will undoubtedly experience a range of speech levels at a range of SNRs in their everyday environments, a range of levels and SNRs should be used to evaluate the hearing aid. Unfortunately, the range appropriate to a particular client is unlikely to be known, and obtaining reliable speech test scores under several measurement conditions is very time consuming. A second potential disadvantage of speech tests is that only the benefit provided by a hearing aid is measured. While this may be appropriate in some circumstances, in others, one may wish to measure the benefit provided by an entire rehabilitation program, including, but not limited to, hearing aid provision. Other benefits to the client may include reduced anxiety as a result of counselling and receiving information about the nature of hearing loss, increased confidence as a result of communication training, and increased use of hearing tactics such as positioning oneself with respect to the talker, use of lipreading, and positioning the talker in the light.

Self-report measures of benefit suffer from neither of these limitations. They have the potential to reflect the clients' assessment of the benefit they have received from rehabilitation, and this benefit may be expressed in terms of decreased disability (e.g., increased ability to detect and identify speech and other sounds) or decreased handicap (e.g., increased participation in social events or decreased negative emotional consequences of a hearing loss). The limitations that self-report measures have depend on how the self-reports are obtained. Self-report methods can be categorized according to how the benefit measure is derived. In the first method, which we will call subtraction of disability or handicap, the clients are asked about how much difficulty they have hearing in a range of situations or what impact these difficulties have on their lifestyle (i.e., handicap), both with and without a hearing aid. In one variation of this method, questions about hearing difficulties without a hearing aid are asked prior to receiving the hearing aid. At some time following hearing aid fitting, the questionnaire is readministered, and the clients answer about the difficulties they now experience. When separately administered, before and after provision of services to clients, such questionnaires have the potential to assess the impact of all components of a rehabilitation program. Any questionnaire that measures hearing difficulties or associated handicap could be used in this way, provided that the inventory has adequate test–retest reliability. One that has frequently been used for this purpose is the Hearing Handicap Inventory for the Elderly (HHIE) (Ventry and Weinstein, 1982; Newman and Weinstein, 1988; Malinoff and Weinstein, 1989a, b). A potential problem with this method is that the benefit of rehabilitation is calculated from the difference between the "before" and "after" scores. The random measurement errors associated with each administration are likely to be independent (because of the time separation between administrations), and so the error variances add. The resulting uncertainty in benefit, for a single client, may be of similar magnitude to the size of the benefit measured. For example, when the HHIE is administered by the clinician, the 95 percent confidence interval for detecting a change is 19 points (Weinstein et al, 1986). The mean change in scores following rehabilitation in several studies is in the range of 15 to 30 points (Chermak & Miller, 1988; Malinoff & Weinstein, 1989b; Mulrow et al, 1992; Taylor, 1993). When the client self-administers the HHIE, the situation is worse, as the 95 percent confidence interval for a change in scores increases to 36 points (Weinstein et al, 1986). Demorest and Walden (1984) contains an excellent treatment of the psychometrics of self-assessment inventories.

In the second variation of the subtractive approach, the before and after measures are both obtained some time after the client has been fitted with a hearing aid. The client is asked to rate separately the difficulty of communicating with and without the hearing aid. The Profile of Hearing Aid Benefit (PHAB) (Cox et al, 1991) was designed to be used in this way, but can also be administered separately before and after rehabilitation. This approach has the advantage that because the aided and unaided response scales lie beside each other, and because the before responses are visible to the client while the client is marking the after responses, the measurement errors are likely to be correlated. Looked at another way, if the client feels that the aid helps a bit, he/she may make sure that a higher rating is given for aided than for unaided listening, even if he/she is not sure what absolute rating should be given for each. The lack of independence is thus a good thing because we are primarily interested in the difference score. When the before score is subtracted from the after score, the resulting difference scores are thus likely to contain less error than if the two scales were completed independently. A potential disadvantage of obtaining both scores after receiving rehabilitation is that clients have to rely on their memory of how much difficulty they had in each situation before they wore a hearing aid, although this is made easier if they still sometimes elect to not wear their hearing aid in each of these situations.

In an alternative to the subtractive approach, this avoidance of two separate sets of measurement errors is carried further. In questionnaires such as the Hearing Aid Performance Inventory (HAPI) (Walden et al, 1984), the client is asked to directly assess the benefit of the hearing aid in each situation, and we will call this the direct differential approach. While the client still has to mentally compare the difficulty in a situation with and without the hearing aid, the client then has to translate this concept of benefit to an appropriate point on only one response scale for each listening situation. This direct differential approach still carries the disadvantage of requiring clients to remember how much difficulty they had without the hearing aid in each situation.

As each of these self-report methods has different advantages and disadvantages, the most efficient methods of quantifying benefit can only be found in studies that quantify the reliability and sensitivity with which each method quantifies the changes in disability or handicap that result from rehabilitation.

All of the above self-report methods have an additional disadvantage in common: the questionnaires consist of a fixed list of listening situations for which the client is required to rate difficulty, handicap, or benefit. For any particular client, many of these situations may be irrelevant because they are not experienced by the client. Even if they are experienced, they may be experienced only infrequently, or they may not be situations in which the client considers ease of listening to be very important. Even if ease of listening is important, the client may already experience little difficulty in some of these situations, so that measurement of hearing aid benefit in the situation is irrelevant.

Irrelevant items can cause at least three problems. First, the answers usually carry equal weight to items that have great importance to the client, and thus decrease the validity of the measure. Second, time is needed to work through the irrelevant items, and this is a problem, particularly if the questionnaire is administered by the clinician. Third, the irrelevance of too many items may cause the client (and/or the clinician) to perceive the whole questionnaire as irrelevant and to cease filling it in (or using it with other clients).

The problem of clients not experiencing situations has been partially addressed by careful selection of questionnaire items. Schum (1992) and Dillon (1994) have both provided abbreviated versions of the HAPI by judicious choices of response items. Cox and Alexander (1995) have similarly provided an abbreviated version of the PHAB. Ventry and Weinstein (1983) have provided an abbreviated version of the HHIE. While this approach minimizes the problem of irrelevance, it does not remove it and does not address the relative importance of different items. Gatehouse (1994) has suggested that when assessing unaided disability, the importance of each situation can be assessed by asking (a) how often the client is in the situation, (b) how difficult the situation is, and (c) how much the difficulty annoys or upsets the client. When aid benefit in each situation is assessed at a later time, he suggests recording (d) the amount of time the aid is used, (e) the degree of help the aid provides, and (f) the client's satisfaction with the aid in this situation. This adds useful information but considerably lengthens the questionnaire. Presumably, the answers to a, b, c, and d for each situation could be combined and used as a weighting factor when determining the overall benefit provided to the client or degree of satisfaction experienced by the client.

A radical alternative is to let each client make up his or her own questionnaire! This solves the problem of irrelevance and minimizes the problem of some items being much more important than others. This method has its origins in the field of mental health (Kiresuk and Sherman, 1968), in which clients were encouraged to form individual "goals" for their behavior. At the conclusion of treatment, clients were then assessed as to how well these goals were attained. The concept of goal planning and assessment was applied to rehabilitative audiology by McKenna (1987). Based on this concept, Dillon et al (1991a, b) implemented two Goal Attainment Scaling (GAS) assessment methods and compared them to various other measures of benefit from rehabilitation. The first of these methods, called specific GAS, required the clients to nominate situations in which they were experiencing difficulty in hearing and in which they would like to improve their ability to hear. At the first interview, their current level of difficulty in hearing in each of these situations was rated on a 5-point scale. In conjunction with the clinician, the clients also established a goal level of hearing ability, in each situation, with which they would be satisfied. (The clinician was involved in negotiating the goal level to ensure that goals were realistic, considering the degree of loss of the client and the acoustic difficulty of each situation nominated.) At the conclusion of rehabilitation (several weeks after aid fitting and again 3 months later), the new level of difficulty of hearing in each situation was assessed on the same scale. Benefit was assessed as the difference between the before and after measures, averaged across the listening situations nominated. Gatehouse (1994) has also used the concept of a questionnaire in which some of the listening situations were nominated by the client. The second implementation of GAS (Dillon et al, 1991b), called global GAS, was similar to the specific GAS method except that all of the items were "standard" common situations. Of all of the methods tried (including the HHIE), the specific GAS method produced the highest correlation with average hearing loss, which was thought to provide a partial validation of the measure. It was also perceived by the clinicians involved in the experiment to be the method most relevant to the clients. As a result, the specific GAS method was adopted throughout the National Acoustic Laboratories' (NAL) hearing centres as the primary method of assessing the outcomes of rehabilitation.

Following a few years of experience with this method, a few problems emerged. Many clinicians did not like negotiating and quantifying goal levels. Some did not like quantifying the degree of difficulty experienced by the client when they first met the client. Statistical analysis of a random sample of 368 clients showed that the high correlation of benefit with hearing loss observed in the experiment did not emerge in routine clinical use (Lovegrove et al, 1992). (The relationship of benefit to hearing loss will be discussed later in this paper.) As a result of these problems, a new but related measure, the Client Oriented Scale of Improvement (COSI), was devised. This paper describes this method and shows that the benefit it measures is at least as reliable and valid as the benefit assessed by a variety of other methods, including the HHIE and a slightly modified version of the PHAB.

When evaluating hearing aid fittings, questionnaire items may use words related to "satisfaction" or words related to "benefit." While the concepts are certainly different, it is difficult to determine to what extent self-report measures of benefit are influenced by the satisfaction clients feel with their hearing aids, or indeed with the way they have been treated by their clinicians. Walden et al (1984) comment that with self-report measures it is common for clients to exaggerate their difficulties when they are first seen and then to exaggerate the improvement at the final evaluation. Presumably, they do the first of these to ensure that their problems are taken seriously and the last to "thank" a clinician or organization that has treated them well. In Dillon et al (1991b), satisfaction with the hearing aid correlated as highly with various measures of benefit as did the other measures of benefit. It correlated more highly with aid use than did any of the benefit measures. To try and assess the impact of treatment satisfaction on self-reported benefit, two measures of treatment satisfaction were included in the current study. In the current study, we will use measures that apparently reflect benefit, measures that apparently reflect satisfaction, and measures that apparently reflect aid use all as potential measures of the "success" of the hearing aid fitting and associated rehabilitation services.

This introduces the problem of how one assesses which of several measures is the best measure. If there was an obviously correct answer for how much benefit or satisfaction a hearing aid provides, and whether clients most value benefit or satisfaction, there would be no need to devise further evaluation methods. The approach used in this paper is a statistical one. All of the measures are initially regarded as potentially useful measures of the success of the hearing aid. A consensus measure is then formed by iteratively removing measures that are not well correlated to the remaining measures and then averaging the remaining measures. We assume that this consensus is the best available estimate of the success of the hearing aid. The accuracy of each method can then be assessed by seeing how well each correlates with the consensus measure.

The primary purpose of this paper was to see which of several methods of assessing self-reported benefit or satisfaction with hearing aids were most suitable for clinical use. We considered that a suitable method would need to have acceptable reliability (as assessed by test–retest stability), acceptable validity (as assessed by its correlation to other measures thought to be valid), be convenient to use, and preferably be capable of improving, as well as just measuring, the rehabilitation outcomes. The first two of these are addressed quantitatively in this paper, and the last two are partially addressed by systematically seeking the views of the clinicians who used each of the methods compared in the experiment. We were particularly interested in the statistical properties of a newly developed method of client assessment, the COSI.

METHOD

Subjects

Subjects were 98 adult clients attending Australian Hearing Services clinics. None of the subjects had worn hearing aids previously. The distribution of three-frequency average hearing level (3FAHL; 500 Hz, 1 kHz, and 2 kHz) in the better ear is shown in Figure 1. The selection criteria were that the subjects be willing to participate in the trial and be able to fill in the HHIE questionnaire (or be capable of answering its questions if someone else could fill in the questionnaire for them). Subjects had an average age of 71 years with an interquartile range from 67 to 75 years. The type of hearing loss was not recorded for analysis, but as the sample is large, the losses should represent the typical Australian Hearing Services new clients caseload: predominantly sensorineural, predominantly presbycusic, predominantly mild to moderate bilateral hearing loss. With the exception of a nominal annual service charge, subjects did not have to pay for their hearing aid(s).

Procedure

Hearing aids and associated counselling were provided in a series of appointments, during and after which various self-report measures were obtained. The data were obtained by 14 audiologists operating from nine clinics. Each audiologist saw from 1 to 13 clients as part of the study.

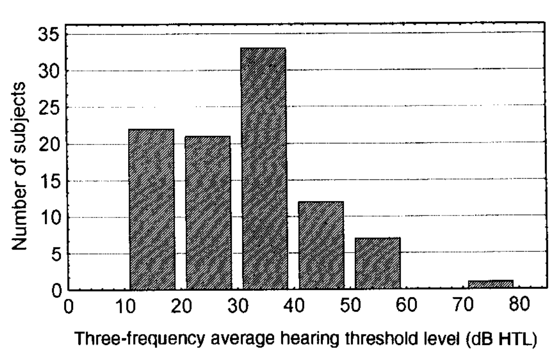

Figure 1 Histogram of 3FAHLs in the better ear of the subjects.

Initial Appointment. Pure-tone audiometry and impedance testing (where necessary) were performed and a history was taken. Hearing aid(s) were prescribed according to the NAL-RP selection procedure (Byrne and Dillon, 1986; Byrne et al, 1990) and ear impressions were taken. During the interview, the first part of the COSI evaluation was completed and at the end of the appointment the HHIE was completed.

Fitting Appointment. The hearing aids fitted (behind the ear or in the ear, as appropriate) were electrically programmable with output-controlled compression limiting and were fitted unilaterally or bilaterally, as appropriate. The frequency response shaping was verified by probe tube measurement of insertion gain, and the subjects were oriented to hearing aid use.

Check-up Appointment. Subject's use and understanding of the hearing aid was checked, problems were dealt with, and subjects were given information on strategies for hearing more effectively in difficult listening situations. This was usually the final appointment within the rehabilitation program but, where necessary, additional check-up appointments were scheduled. The last check-up appointment will be referred to throughout this paper as the exit appointment. During the exit appointment, the second part of the COSI assessment was administered, and at the end of the exit appointment the HHIE was readministered. The median time from fitting to the exit appointment was 5.7 weeks, with an interquartile range from 4.4 to 7.3 weeks.

Follow-up Appointment. Subjects were recalled, nominally 3 months after fitting. In fact, the median time between fitting and follow-up was 16 weeks, with an interquartile range from 12 to 22 weeks. A range of self-assessment reports, as described below, was administered during and at the end of the 3-month follow-up appointment.

Self-Report Measures

Client Oriented Scale of Improvement (COSI). As outlined in the introduction, this measurement was a development of the GAS method reported in Dillon et al (1991b). The form used for recording the data is shown in the Appendix. Administration commenced at the test appointment, at which time the audiologist asked the client to nominate up to five specific situations in which the client would like to cope better. These situations, described as specifically as possible, were recorded and filed. Subjects were (mostly) also asked to state which of these situations was most important, which was second in importance, and so on. At the exit appointment, each of the descriptions was read back to the client, and for each situation the client was asked (a) how much better (or worse) the client could now hear (i.e., ease of communication after rehabilitation relative to before rehabilitation), and (b) how ably the client could now hear (i.e., absolute ease of communication after rehabilitation). The five response choices available to the client are those shown on the form in the Appendix. For scaling purposes, the response choices were assigned the numbers 1 to 5, with 5 corresponding to "much better" and "almost always" for questions a and b, respectively. Both questions were repeated at the follow-up appointment. Subjects were not advised of the answers that they had previously given. Four different summary measures were obtained from each subject's data. The first of these was the average improvement at the exit appointment, obtained by averaging, across all of the listening conditions, the scores for question a. This will be referred to as "improvement." The second measure was similarly obtained by averaging the scores for question b and will be referred to as "final ability." The third and fourth measures were obtained by applying the same processes to the data obtained at the 3-month follow-up, rather than at the exit appointment.

Hearing Handicap Inventory for the Elderly (HHIE). The HHIE questionnaire (Ventry and Weinstein, 1982), comprising 25 items, was modified slightly, as specified in Dillon et al (1991a). The first modification, to encourage more responses at either end of the scale, was the replacement of the "yes" option by "yes or usually" and replacement of the "no" option by "no or rarely." The second modification was a minor rewording of three questions, to make these questions more applicable to the lifestyle of many elderly Australians. The questionnaire items were scaled so that a score of 0 represented most handicap and a score of 4 represented no handicap. (This is the reverse of how HHIE scores are usually reported, but throughout this paper, all self-report scores are organized so that higher numbers represent the "best" hearing for the subject: minimum handicap or maximum benefit from rehabilitation, as appropriate.) The HHIE items were also divided into a social/situational subscale of 12 items and an emotional subscale of 13 items. The first of these reflects the impact of hearing loss on the activities in which the client engages, while the second reflects how the hearing loss causes the client to feel. The HHIE was administered at the end of the initial appointment, the end of the exit appointment, and the end of the follow-up. The first few questions were administered by the clinician, and the next few questions were self-administered, with supervision by the clinician, after which the remainder of the questions were self-administered. We will refer to this as "supervised self-administration." The degree of supervision was generally less for the re-administrations than for the initial one. For a few less able clients, all questions were administered by the clinician.

Modified Profile of Hearing Aid Benefit (MPHAB). After some pilot data using the standard PHAB questionnaire (Cox et al, 1991) were collected, the questionnaire was modified to simplify the task. The first modification was the deletion of items 13, 29, 35, 36, and 53, as these were considered to have low applicability to elderly Australians; item 5 was deleted by accident. The second modification was the replacement of some words by alternatives more commonly used in Australia. The third modification was motivated by the observation that some clients, for some questions, were not able to correctly interpret whether the "always" end of the scale meant that communication was always easy or always difficult. Similar difficulties with the PHAB by some subjects were mentioned by Cox and Rivera (1992). To avoid this difficulty, all questions relating to ease of understanding speech were re-expressed to be positive statements, so that the "always" end of the 7-point scale consistently corresponded to easy communication. This meant that subjects only had to understand the situation being described, not the syntax of whether ease or difficulty of communication was being described. For most items, the modification could be made by changing one or two words (e.g., "can understand" for "have trouble understanding"). For a few items, more extensive rewording was needed. Although we do not believe that these changes will materially alter the aspects of benefit assessed by the questionnaire, nor decrease its reliability, we will refer to the altered questionnaire as the MPHAB. Responses to the seven response categories were assigned the numbers 1 to 7, with 1 representing "never intelligible" or "always uncomfortable" and 7 representing "always intelligible" or "never uncomfortable," as appropriate. Of the 60 items in the MPHAB questionnaire, 45 related to ease of intelligibility and 15 related to loudness discomfort. Two subscale averages were formed from these items; these will be referred to as "MPHAB intelligibility" and "MPHAB comfort," respectively. The MPHAB was completed by supervised self-administration at the end of the 3-month follow-up. A few clients found self-administration sufficiently difficult that they needed the audiologist to administer the entire questionnaire.

Shortened Hearing Aid Performance Inventory for the Elderly (SHAPIE). The SHAPIE questionnaire (Dillon, 1994) is a shortened, 25-item version of the HAPI (Walden et al, 1984). The shortening was carried out in such a way as to maximize the statistical power of the items by retaining items that had the best combination of applicability, high intersubject standard deviation, high item-total correlation, and high factor loading. The SHAPIE items ask about the helpfulness of a hearing aid in various listening situations. Scores of 1 to 5 were assigned to the responses, with 1 corresponding to "hinders" and 5 corresponding to "very helpful." The SHAPIE questionnaire was administered by the audiologist at the end of the 3-month follow-up for 42 subjects and self-administered (supervised) for the other 56 subjects.

Hearing Aid Users' Questionnaire (HAUQ). The HAUQ (Forster and Tomlin, 1988) contains a range of questions relating to hearing aid use. Five of these provide data that may be related to hearing aid benefit and satisfaction. Question 2 asks how much the client uses the hearing aid, with responses of "never," "less than 1 hour per week," "less than 1 hour per day," "1 to 4 hours per day," "4 to 8 hours per day," and "8 or more hours per day." These were assigned scores of 1 to 6 respectively. Question 3 asks how much help the hearing aid provides in each of six generalized listening situations (family, small group conversation, meetings, social activities, television/radio, and telephone). Response choices, on a 3-point scale, are "a lot," "a little," and "not at all." These were assigned scores of 1, 2, and 3, respectively. A fourth response of "help not needed" was not assigned any score. Question 4 asks whether the client has difficulties with each of six potential hearing aid problems: positioning/removing the hearing aid, manipulating controls, feedback whistling, earmold/shell discomfort, loudness discomfort, and own voice quality. For each of these, a yes or no answer is required, and these were scaled as 0 or 1, respectively. Question 5 asks how satisfied the client was with the hearing aid. The four response alternatives are "very dissatisfied," "dissatisfied," "satisfied," and "very satisfied." These were assigned scores of 1 to 4. Finally, question 7 asks how satisfied the client was with the service received, with the same four response alternatives. Each of the HAUQ questions 1 to 7 was treated as a separate subscale, with the number of items in each subscale varying from one to six. The HAUQ was administered by the audiologist at the 3-month follow-up.

Satisfaction Scales. Because satisfaction, measured with question 5 of the HAUQ, had previously been shown to be correlated with many other benefit measures, two additional overall satisfaction questions, with a more finely graduated response scale, were devised and included. In the first of these, subjects were asked to mark, on a 0 to 100 scale, how satisfied they were overall with their hearing aids. They were advised that 0 would mean that they were not at all satisfied and 100 would mean that they were totally satisfied. The second item was similar, except that they were asked how satisfied they were overall with the help they received from Australian Hearing Services with their hearing problems. These questions will be referred to as the Aid Satisfaction Scale (ASS) and Service Satisfaction Scale (SSS), respectively. These scales were self-administered following the 3-month follow-up.

Data Analysis

The data were analyzed in four ways. First, summary statistics for each measure were prepared. Second, the correlations between the measures obtained at the 3-month follow-up were shown with a correlation matrix. The third analysis was aimed at determining which measures were best correlated with the overall best estimate of success. To do this, scores for each scale or subscale were standardized so that they had zero mean and unity standard deviation. These standardized scores were then regarded as "items" in a questionnaire, and an item-total correlation analysis was performed. In this statistic, the correlation is found between each item and the sum of all of the other items. Items with a low item-total correlation thus measure something different from the rest of the items. In the fourth analysis, data obtained at the 3-month follow-up were compared with data obtained at exit from the last regular appointment. Mean changes, test–retest standard deviations, and correlations were examined.

Clinicians' Opinions

Because the clinician is an important link in the outcomes assessment chain, the audiologists participating in the trial were asked to evaluate each measure. After the data were collected, the audiologists were provided with a questionnaire that asked them to separately rate the usefulness to the clinician and administration convenience of each measure. The clinicians were advised that because each measure could be administered by the clinician or self-administered by the client, they should answer assuming the method of administration that they consider to be most suitable. Responses were allowed on a 4-point scale with 4 representing definitely useful (or convenient) and 1 representing definitely not useful (or convenient). The clinicians were not aware of the statistics of any of the measures when they were surveyed. Results were analyzed only for 12 of the 14 audiologists as the other 2 participated in the development of one of the measures.

RESULTS

Mean and Range of Data

Mean scores, standard deviations, medians, and 25th and 75th percentiles are shown for each measure obtained at the initial and exit appointments in Table 1. The same descriptive statistics for the measures obtained at the 3-month follow-up are shown in Table 2. Most of the measures provided unremarkable data. All benefit measures (SHAPIE, HHIE, MPHAB intelligibility, COSI improvement, and HAUQ3 aid benefit) showed that the clients, on average, experienced considerable benefit from the hearing aid. The MPHAB comfort scores showed that clients experienced more loudness discomfort when they were wearing a hearing aid than when they were unaided, which is also unremarkable. The satisfaction scores (HAUQ5, HAUQ7, ASS, and SSS) were very high. The two measures related to satisfaction with service (HAUQ7 and SSS) had extremely high scores and showed little dispersion between subjects.

Relations between Measures at 3-Month Follow-up

Table 3 shows all of the significant correlation coefficients between the measures obtained at the 3-month follow-up. All coefficients involving the measures HAUQ2, HAUQ4, HAUQ5, and HAUQ7 are Spearman's rho (because the measures are rank order statistics, without significant averaging across items). All other coefficients are Pearson's product-moment correlation coefficients. (Correlation coefficients and significance levels calculated by the two methods were, in fact, extremely similar.) As would be expected, most of the coefficients are positive. The only exceptions are correlations between the MPHAB comfort scores and the other scores, which means that those who reported the greatest benefit from their hearing aids also reported the greatest increase in loudness discomfort.

As described in the Method section, agreement between the various measures was quantified using item-total correlation. After the scores for each of the 14 measures shown in the left-hand column of Table 4 were standardized, item-total correlations were found; these are shown in the next column. Clearly, the MPHAB

Table 1 Descriptive Statistics for Measures Obtained at the Initial Appointment and at the Exit Appointment

Questionnaire	*Mean*	*Median*	*SD*	*25th Percentile*	*75th Percentile*
HHIE at entry	2.07	2.00	0.86	1.44	2.80
HHIE at exit appointment	3.32	3.60	0.68	2.96	3.92
HHIE change	1.26	1.20	0.90	0.48	1.92
COSI improvement	4.29	4.40	0.64	4.00	4.80
COSI final ability	4.19	4.33	0.63	3.75	4.66

Table 2 Descriptive Statistics for Each of the Outcome Measures Obtained at 3-Month Follow-up

Questionnaire	*Mean*	*Median*	*SD*	*25th Percentile*	*75th Percentile*
SHAPIE	4.07	4.00	0.46	3.80	4.36
HHIE at 3 mo	3.32	3.44	0.64	3.04	3.84
HHIE change	1.25	1.08	0.89	0.64	1.84
MPHAB intelligibility, aided	5.77	5.89	0.82	5.36	6.38
MPHAB intelligibility, unaided	3.69	3.73	0.97	3.00	4.38
MPHAB intelligibility, change	2.08	2.16	1.19	1.35	2.82
MPHAB comfort, aided	4.43	4.53	1.20	3.80	5.27
MPHAB comfort, unaided	5.68	5.80	0.78	5.27	6.20
MPHAB comfort, change	−1.25	−1.13	0.98	−1.80	−0.53
COSI improvement	4.38	4.50	0.54	4.00	4.80
COSI final ability	4.33	4.40	0.59	4.00	4.80
HAUQ2 (aid use)	4.92	5.00	0.71	4.00	5.00
HAUQ3 (aid benefit)	2.76	2.81	0.29	2.67	3.00
HAUQ4 (aid problems)	0.90	1.00	0.14	0.83	1.00
HAUQ5 (aid satisfaction)	2.64	3.00	0.50	2.00	3.00
HAUQ7 (service satisfaction)	2.95	3.00	0.22	3.00	3.00
ASS (aid satisfaction)	87.2	90	14.3	80	100
SSS (service satisfaction)	97.8	100	4.6	98	100

Values are points assigned to the response scales, as specified in the Method section. The "change" scores (for HHIE and PHAB questionnaires) were obtained by subtracting the unaided scores from the aided scores.

comfort measure, with the least positive correlation coefficient, is measuring something very different from the remainder of the measures. Consequently, it was deleted and the item-total correlation repeated with the remaining 13 items. This process of deleting one item each iteration was repeated until there were only three items left. Detailed results for the first five steps are shown in Table 4. Items are listed so that at step 5, their item-total correlations are in descending order. The bottom two rows show the summary statistics for the items remaining at each step. Cronbach's alpha describes the average split-half test reliability coefficient for the

Table 3 Correlation Coefficients between Measures Obtained at the 3-Month Follow-up

	SHAPIE	*MPHAB Intelligibility*	*MPHAB Comfort*	*HHIE Social Situational*	*HHIE Emotional*	*COSI Improvement*	*COSI Final Ability*	*ASS*	*SSS*	*HAUQ 2*	*HAUQ 3*	*HAUQ 5*	*HAUQ 7*
SHAPIE	—	0.42		0.38	0.29	0.55	0.48	0.52			0.61	0.38	
MPHAB intelligibility	0.42	—	0.47	0.39	0.36			0.36		0.37	0.28	0.35	0.31
MPHAB comfort		−0.47	—		−0.29								
HHIE social/situational	0.38	0.39		—	0.81	0.27					0.31		
HHIE emotional	0.29	0.36	−0.29	0.81	—								
COSI improvement	0.55			0.27		—	0.66	0.43		0.27	0.51		
COSI final ability	0.48					0.66	—	0.41			0.39		
ASS	0.52	0.36				0.43	0.41	—	0.36	0.31	0.41	0.53	
SSS								0.36	—				0.32
HAUQ2		0.37				0.27		0.31		—	0.27	0.28	
HAUQ3	0.61	0.28		0.31		0.51	0.39	0.41		0.27	—	0.38	
HAUQ5	0.38	0.35						0.53		0.28		—	0.31
HAUQ7		0.31							0.32			0.31	—

The MPHAB and HHIE scores used were difference scores (aided minus unaided) with the unaided score obtained at the initial appointment for the HHIE. Only correlations significant at $p < .01$ are shown. HAUQ4 is not shown as it was not significantly correlated to any of the other measures.

Table 4 Item-Total Correlations Resulting from the First Five Steps of the Iterative Culling Process

	Step				
Measure	*1*	*2*	*3*	*4*	*5*
SHAPIE	0.69	0.68	0.67	0.68	0.70
HAUQ3 (aid benefit)	0.59	0.58	0.60	0.62	0.61
ASS (aid satisfaction)	0.60	0.61	0.62	0.59	0.60
COSI improvement	0.59	0.57	0.56	0.57	0.59
MPHAB intelligibility	0.41	0.48	0.51	0.52	0.51
HHIE social/situational	0.46	0.48	0.49	0.49	0.50
COSI final ability	0.52	0.51	0.50	0.50	0.49
HAUQ5 (aid satisfaction)	0.51	0.52	0.50	0.50	0.48
HHIE emotional	0.33	0.36	0.38	0.38	0.38
HAUQ2 (aid usage)	0.33	0.35	0.37	0.38	0.36
HAUQ7 (service satisfaction)	0.33	0.34	0.36	0.34	
SSS (service satisfaction)	0.29	0.32	0.31		
HAUQ4 (aid problems)	0.01	−0.03			
MPHAB comfort	−0.27				
Number of items	14	13	12	11	10
Cronbach's alpha	0.763	0.806	0.831	0.833	0.833
Average interitem correlation	0.20	0.23	0.30	0.32	0.35

composite measure. High values (close to 1.00) are obtained if the composite measure contains many items or if the items are highly correlated with each other. As can be seen, Cronbach's alpha increases with the first two deletions, and then remains relatively stable with subsequent deletions. (Presumably, each subsequent decrease in the number of items was approximately compensated by the increase in the correlations between the remaining items.) The average interitem correlations are shown in the last row of the table. Figure 2 shows how Cronbach's alpha, and the average interitem correlation, varied over all 12 iterative steps.

At what point should the culling process be terminated? If the criterion is to maximize Cronbach's alpha, then anywhere from 2 to 12 deletions would produce similar results. We elected to delete four measures. The first two were deleted because their item-total correlations were much lower than those of the remaining items. The items with the next two lowest item-total correlations were deleted partly on the basis of their content: both were measures of satisfaction with the quality of service received, whereas all of the remaining items were centered on the hearing aid, or on reduction of disability or handicap. Each measure remaining therefore agrees with the average of the remaining measures to the extent shown in the step 5 column of Table 4. (In fact, the rank order of the measures does not change much if a fewer or greater number of measures has been retained for the consensus measure.)

An additional analysis was performed to investigate the effect of alternative scoring methods for the MPHAB and COSI measures. Only those items in the MPHAB that appear in the

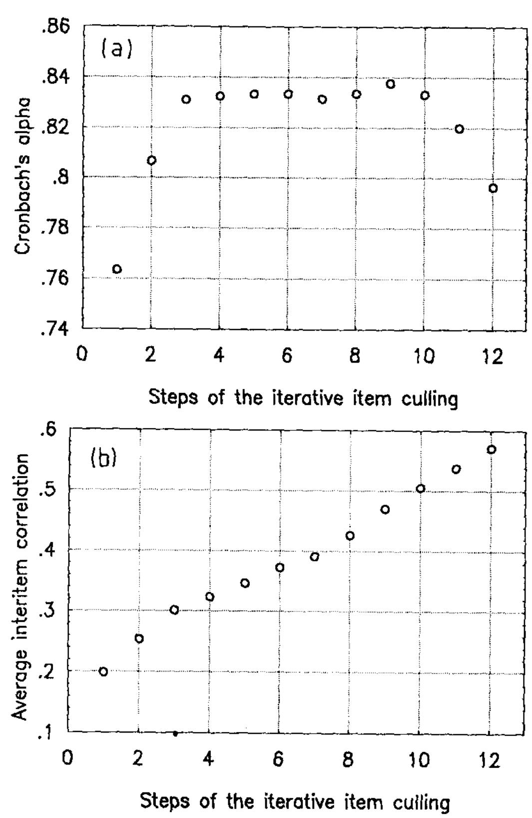

Figure 2 Cronbach's alpha (a) and average item-total coefficient (b) for each step of the iterative item-total correlation culling process.

APHAB were scored (Cox, personal communication). This resulted in 6 items for the comfort subscale and 16 items for the intelligibility subscale. For the COSI data, the priorities assigned by the client were used as weighting factors when the data were averaged across listening situations. The highest priority situation received a weighting of 5, the next a weighting of 4, and so on. Priority data were only collected for 79 of the clients. For these clients, changes in the item-total correlations caused by the alternative scoring methods at step 5 of the item-total analysis were less than 0.01 for all measures, except for the MPHAB intelligibility coefficient, which decreased by 0.04, and the COSI final ability coefficient, which increased by 0.06. All changes resulting from this alternative scoring were thus small.

Test–Retest Differences and Reliability

Table 5 shows test–retest correlation coefficients and the mean and standard deviation of test–retest differences for the two measures that were measured at both the exit and 3-month follow-ups. On the basis of the test–retest correlation coefficients, the COSI scores appear to be more reliable than the HHIE scores. Although the standard deviations of the test–retest differences are also smaller for the COSI than for the HHIE (apparently indicating better reliability for the COSI), Table 2 shows that the range of scores across subjects is also smaller for the COSI than for the HHIE, making it difficult to compare the two measures in this way.

An additional analysis was performed to see if the benefit of rehabilitation that was reported 3 months after aid fitting could be adequately estimated at the time of the exit appointment. To test this, a global measure of benefit/satisfaction was formed by summing the standardized scores from the items with the highest item-total correlations shown in step 5 of Table 4, but excluding the two measures that were available at both appointments. Thus, questionnaires included in the global measure were SHAPIE, HAUQ3, ASS, MPHAB intelligibility, and HAUQ5. Correlations between the two COSI measures and the two HHIE measures, obtained at each appointment, with the global measure obtained at the 3-month follow-up, are shown in Table 6. The 3-month global measure could be predicted from the COSI and HHIE scores at the exit appointment at least as well as from the COSI and HHIE scores at the 3-month follow-up. Although all of the correlations are significant ($p < .005$), none of the differences in correlation between the two assessment times are significant at $p = .05$.

Factor Analysis

A principal components factor analysis was carried out using the same 14 scores shown in Table 4 as the variables. Factor loadings (i.e., correlations between each variable and each of four factors) after varimax factor rotation are shown in Table 7. The measures loading highly onto factor 1 are the direct differential estimates of aid benefit (COSI improvement, SHAPIE, and HAUQ3), the measure of final communication ability (COSI final ability), and the two measures of aid satisfaction. The only measure loading onto factor 2 are the two HHIE subscales. Factor 3 contains a mixture of aid usage, MPHAB intelligibility improvement, and negative loadings for aid comfort and aid problems. The only measure with a high loading on factor 4 is satisfaction with service, although the HAUQ7 measure of the same concept also has a moderate loading onto this factor. (A high loading for HAUQ7 on any factor is unlikely because of the large number of ceiling scores obtained with the HAUQ7 question.)

Table 5 Test–Retest Statistics for the Measures Administered Twice

	Correlation Coefficient	*Mean Difference*	*SD*
HHIE social/ situation	0.58	−0.03	0.66
HHIE emotional	0.54	0.02	0.63
HHIE total	0.57	−0.01	0.61
COSI improvement	0.73	0.09	0.45
COSI final ability	0.84	0.15	0.35

The mean difference is the 3-month score minus the exit appointment score.

Table 6 Correlation Coefficients between Global Measure of Benefit/Satisfaction at 3-Months Postfitting and Each of the Measures Shown, at Exit Appointment and 3 Months

	Exit Appointment	*3-Month Follow-up*
COSI improvement	0.61	0.54
COSI final	0.53	0.47
HHIE social/situational	0.43	0.39
HHIE emotional	0.31	0.36
HHIE total	0.40	0.36

Table 7 Factor Loadings of Each Variable

	Factor			
Measure	*1*	*2*	*3*	*4*
COSI improvement	**0.80**	−0.18	−0.10	−0.02
SHAPIE	**0.76**	−0.30	0.05	−0.10
HAUQ3 (benefit)	**0.75**	−0.12	0.26	0.05
COSI final ability	**0.72**	−0.13	−0.12	−0.12
ASS (aid satisfaction)	**0.63**	−0.01	0.24	−0.42
HAUQ5 (aid satisfaction)	**0.55**	0.12	0.30	−0.37
HHIE emotional	0.06	**−0.92**	0.09	−0.07
HHIE social/situational	0.24	**−0.88**	0.06	−0.04
HAUQ4 (aid problems)	0.16	0.15	**−0.66**	−0.32
MPHAB intelligibility	0.25	−0.39	**0.60**	−0.24
HAUQ2 (aid usage)	0.35	0.03	**0.58**	−0.02
MPHAB comfort	0.14	0.29	**−0.49**	0.48
SSS (service satisfaction)	0.12	−0.12	−0.08	**−0.83**
HAUQ7 (service satisfaction)	0.19	−0.03	0.41	**−0.41**

Factors are listed in order of decreasing loading against their highest loading factor (shown in bold).

Audiologists' Opinions

The results of the opinion surveys obtained from 12 of the 14 participating audiologists are shown in Table 8. Friedman's analysis of variance by ranks indicated that there were significant differences among the ratings of usefulness ($p < .00001$) and convenience ($p < .00001$). Inspection of the medians indicate that this is principally due to low ratings for the usefulness of the satisfaction scales and the SHAPIE and MPHAB questionnaires and to low ratings for the convenience of the MPHAB questionnaire and, to a lesser extent, the HHIE, SHAPIE, and satisfaction questionnaires.

DISCUSSION

If one were to choose an assessment measure based solely on the extent to which each measure agrees with the consensus measure, then Table 4 indicates that SHAPIE would be chosen. It is reassuring that the SHAPIE measure proved to be among the most accurate, as the direct differential test format and the basis upon which items were chosen to abbreviate the original HAPI into the SHAPIE both should help achieve accurate results. While it may be a good assessment measure to use in research studies, we do not, however, consider it the best for routine clinical use. Like all assessment measures based on questionnaires, it contains items that are unimportant for any particular client. Furthermore, it is time consuming if administered by the clinician. Even when self-administered, clinical time is required for scoring and interpretation. Although the score (and subscale scores if desired) can be easily interpreted as a reduction in disability across the situations covered by the scale or subscale, there is no direct connection between these scores and any further remedial action required of the clinician. We presume that reasons of this kind lay behind the relatively low convenience and usefulness ratings given by the clinicians, as shown in Table 8. For similar reasons, we do not consider the MPHAB or HHIE measures to be optimal for clinical use. In addition, both of these questionnaires were significantly ($p < .05$) less correlated with the consensus measure than was the SHAPIE questionnaire. It is possible that the clinicians in this trial may have found the

Table 8 Median Audiologists' Opinions about Usefulness and Convenience of Each Measure

Measure	*Usefulness*	*Convenience*
HHIE	3.5	3.0
SHAPIE	3.0	3.0
MPHAB	3.0	1.0
COSI improvement	4.0	4.0
COSI final ability	4.0	4.0
Satisfaction scales	2.0	3.0
HAUQ2 (usage)	3.5	4.0
HAUQ3 (benefit)	4.0	4.0
HAUQ4 (problems)	4.0	4.0
HAUQ5 (aid satisfaction)	4.0	4.0
HAUQ7 (service satisfaction)	4.0	4.0

SHAPIE, MPHAB, and HHIE questionnaires more useful if they had been provided with information about how to obtain subscale scores. In particular, scores on the loudness aversiveness subscale of the MPHAB can directly suggest appropriate clinical action (reduction of hearing aid maximum output). It is likely that use of the shorter APHAB (which was not available when this study commenced) would have improved clinicians' estimates of convenience. However, its length is similar to that of the SHAPIE (24 vs 25 questions), and it has to be completed separately for unaided and aided listening, so it is unlikely to be rated as any more convenient than the SHAPIE.

The second measure in Table 4 (HAUQ question 3) has a reasonable item-total correlation (i.e., correlation with the consensus measure) and, with only six items, is much more convenient than the longer questionnaires. On the basis of this study, it appears to be a reasonable outcomes measure, though for the reasons listed below, it is not our most highly recommended one.

Although the third measure (ASS) listed in Table 4 is very different from these questionnaires, it has some disadvantages in common, and we also recommend against its routine use. While the simple question about the client's satisfaction with the hearing aid may provide an assessment that is as reliable as that obtained by more time-consuming techniques, the single number obtained is not very informative. Even though a high number may indicate that the rehabilitation program can be terminated with an acceptable outcome, a low number does not indicate why the client is dissatisfied, what the nature of the remaining problems is, or give any clues as to what the clinician should do next.

By contrast, the measure listed as fourth in Table 4, the COSI measure of communication improvement, suffers from neither excessive length, irrelevance, nor lack of diagnostic utility. We recommend it for routine clinical use for the following five reasons, only the last three of which have been examined directly in this experiment. First, and most importantly, the method fits into a well-conducted clinical interview in an unintrusive manner. In essence, the method consists of finding out (a) what the clients say their problems are, (b) how much those problems have been alleviated by the rehabilitation process, and (c) how well the clients are now coping in the situations that were important to them. As these are steps that an effective clinician would normally do, the only extra work that the method requires of the clinician is to systematically document item a and then quantify items b and c.

Second, the client's responses are more likely to be diagnostically useful to the clinician than would be the responses of a much longer questionnaire. For each of a small number of situations, the clinician can consider whether additional action is likely to result in improved outcomes. This additional action might consist of providing assistive listening devices, teaching hearing strategies, providing communication training, or modifying the hearing aid amplification characteristics. While it is possible to do any of these things in the absence of an outcomes assessment measure, time constraints inevitably mean that they cannot all be done. Assessment of outcomes in situations important to the client thus allow programs to be more appropriately individualized. As well as the clinicians obtaining useful information, the clients view the measure as being relevant to and showing interest in them, rather than as one of the anonymous feedback questionnaires that are becoming increasingly common in commercial and public organizations.

Third, the correlation between the COSI improvement measure and the consensus measure is reasonable. The best estimate available is that it ranks fourth in the list of 14 measures. Statistically, its item-total correlation of 0.59 (Table 4, step 5) is not significantly lower ($p = .20$, Fisher's r to z transform) than that of the highest ranking measure (SHAPIE).

Fourth, its test–retest reliability is reasonable. It appears to be as reliable as the HHIE (see Table 5) even though it is based on only three to five items. It certainly seems unusual that a very short questionnaire can be as reliable as a much longer questionnaire. Possibly, people are able to give more precise information when they are being asked about situations that are important to them than when they are being asked about more general situations, some of which have no importance to them. It is also possible that the subjects remembered their earlier responses when they were retested but, given the intervening time and the number of questionnaires administered to the subjects, this seems unlikely.

Last, and presumably for the above reasons, clinicians rate the COSI method as being useful and convenient for them to use. There are three other potential advantages that we have become aware of since conducting the study, but these have in no way been established in the experiment. First, establishing that there are situations in which the client would like to hear better ensures that the client acknowledges that a disability exists. Second, clearly documenting the individual listening problems helps provide case continuity if the clinician changes during the rehabilitation program. Third, listening situations considered important by clients can be categorized, quantified, and prioritized. This can help an organization direct its development or services to the areas of greatest need. All of these things can be achieved without COSI, but all of them follow naturally from its use.

The major disadvantage of the COSI method is that interpreting results accumulated across clients (e.g., to produce composite results for a rehabilitation organization) is less straightforward than for traditional questionnaires. For example, the mean score of 4.38 shown in Table 2 corresponds to a degree of improvement approximately one third of the way from "better" to "much better." However, this degree of improvement is an average across a diverse range of listening situations. For this reason, the COSI form shown in the Appendix contains an additional column labelled "category" and a listing of 15 standard categories that look more like situations described in traditional questionnaires. A more interpretable degree of improvement, accumulated across clients, thus requires each of the individual listening needs to be appropriately categorized, and then for results to be averaged separately for each listening category.

Weighting of the COSI improvement scores by the priority assigned to each listening situation by the subjects neither increased nor decreased its correlations with the other measures. As it complicates the scoring procedure, there is no reason to use the priority information when calculating the summary improvement score for each subject. The subject's priorities may provide some clinically useful information, but the COSI method can be used without ascertaining what priority the client places on each situation nominated.

As well as noting which measures correlate highly with the consensus measure of benefit/satisfaction, it is instructive to note which measures had a low correlation. First, as can be seen from the correlation matrix (see Table 3) or from its place in the item-total correlation (see Table 4), aid usage was poorly correlated with the other measures. Because there were reasonable correlations between the other measures, we cannot ascribe this to the benefit and satisfaction measures being unreliable or insensitive. Instead, we must either view the aid usage data as being unreliable or accept the implication that aid usage and aid benefit/satisfaction are inherently different concepts. If the latter, then we cannot judge the success of a rehabilitation program by how much the hearing aid is worn. (In the extreme case, if the aid is never or very rarely worn, we probably can reach conclusions about the lack of success of a hearing aid fitting.) Cox et al (1991) reviewed several studies indicating that aid use is poorly related to benefit.

Second, average hearing loss in the better ear is not correlated to aid benefit or satisfaction. While it is possible that other combinations of pure-tone loss (e.g.,, four-frequency average loss in the aided ear) may be related, this seems unlikely. Again, the many significant correlations between the various benefit/satisfaction measures means that we cannot explain this finding away by saying that these other measures are insensitive. This lack of correlation should be interpreted in light of the range of hearing loss among the subjects. All but one had 3FAHLs in the better ear in the range 10 to 60 dB HL, and 79 percent of the clients had 3FAHLs better than 40 dB HL. One would expect that significant correlations with benefit must emerge if the range was extended downward to include people with normal hearing in both ears and upward to include people with severe hearing loss in both ears! For hearing losses within the range tested in this experiment, however, the use of average hearing loss as a predictor of likely benefit appears to be totally without support.

Third, the absence of problems with the hearing aid was not positively correlated with benefit/satisfaction. Question 4 of the HAUQ showed no significant correlations with any of the other measures. The loudness aversion (i.e., comfort) items of the MPHAB showed a negative correlation with benefit. That is, those who showed the most benefit tended to report the greatest increase in loudness discomfort with the hearing aid. These two measures were the first two removed in the iterative process of arriving at a consensus measure. The absence of problems with a hearing aid fitting should thus not be regarded as an indication of a successful hearing aid fitting. This does not mean that problems with a hearing aid can be ignored. We certainly see a useful clinical role for questionnaires like HAUQ question 4.

The lack of change of self-reported benefit/satisfaction with time after the aid fitting may be surprising if one believes that new aid users take a considerable time to get used to their hearing aids. Gatehouse (1993) has shown that objectively measured benefit can increase with time after fitting. The lack of change of self-reported benefit in this experiment, however, is evident in several ways. The mean reported benefit at the exit appointment is almost identical to that at 3 months postfitting (see Table 5). Similar results, on the basis of HHIE scores, were reported by Dillon et al (1991b). Of course, no change in average benefit with time might be caused by some subjects reporting more benefit with time and others reporting less. However, the correlation between the self-reported benefit at the two times is high, at least for the two COSI measures, so this cannot have happened. Even more surprisingly, the global measure of benefit/satisfaction obtained 3 months after fitting is just as highly correlated with the COSI scores and HHIE scores at the exit appointment as it is with these scores obtained at the 3-month follow-up (see Table 6). This suggests that it is not necessary to wait more than 6 weeks after fitting before obtaining self-report assessments of the benefit provided by the hearing aid. This is consistent with the data of Mulrow et al (1990), who showed that the self-assessed benefit at 6 weeks after fitting was maintained at much longer intervals. By contrast, both Taylor (1993) and Malinoff and Weinstein (1989b) have shown that, 3 weeks after fitting, the self-assessed benefit is higher than when reassessed at later times.

Certain aspects of this experiment were not as tightly controlled as is common in experiments. For example, time from fitting to outcome measurement varied somewhat, and we do not know how extensively each clinician trained and supervised the subjects as the subjects commenced self-administering several of the questionnaires. Loose control was an intentional part of the design philosophy from the outset, as we wished to determine how the various measures would function in real-world circumstances, rather than when administered by a few clinicians who had a special interest in one or more of the measures.

It is unfortunate that we were not able to investigate more thoroughly whether client's answers to questionnaires on benefit are affected by how well they feel that they have been treated by the service provider. We had anticipated that the two measures of satisfaction with service (SSS and HAUQ7) would allow us to control for this effect in some way. The data for these two measures had a strong ceiling effect (see Table 2). While this may be gratifying for the service providers, it is difficult to use this to make any strong tests of whether reported benefit is independent of satisfaction with service. There is some evidence that this is the case. In Table 3, the only significant correlations involving the two service satisfaction measures are with other satisfaction measures. The ceiling effects were thus not strong enough to prevent these two measures from correlating with anything else. By contrast, the aid satisfaction measures (ASS and HAUQ5) both correlated with numerous measures of aid benefit. It is thus possible that subjects can report benefit without being too influenced by how well they were treated. A more definite test of this would require rewording of the service satisfaction questions to prevent so many responses at the upper end of the scale.

Throughout this study, we have not assumed that benefit (defined as a reduction in handicap or as a reduction in disability), satisfaction, or usage is the more important attribute to measure when measuring the success of rehabilitation. Certainly, the correlations between benefit and satisfaction are much higher than those between usage and the other two attributes, as discussed previously. Separating benefit from satisfaction may not be so easy. After reviewing several studies, Cox et al (1991) concluded that speech understanding (i.e., an objective measure of disability reduction) accounts for less than 40 percent of the variance in users' satisfaction with a hearing aid. Similarly, in the current study, the highest correlation coefficient between self-reported benefit and satisfaction was 0.52, which accounts for only 27 percent of the variance. However, the highest correlation between any two benefit measures was only 0.61, and the highest correlation between any two satisfaction measures was only 0.53, so the low correlation coefficients may be reflecting measurement error rather than the independence of the quantities being assessed. The commonality between satisfaction and benefit is also reflected in the factor analysis results (see Table 7). When four factors were used, although service satisfaction loaded onto its own factor, the two aid satisfaction measures loaded onto the same factor as four of the benefit measures. We could not find any number of factors that resulted in separation of measures associated with satisfaction from those of benefit, and in which tests purportedly measuring the same construct loaded highly onto the same factor. Not surprisingly, as the number of factors increase, factors could increasingly be labelled as particular tests. With six factors, for example, the factors could be labelled as benefit, HHIE, MPHAB, service satisfaction, aid satisfaction and usage, and aid problems. As it is not clear why factors should be captured by particular tests, we cannot interpret these results with any confidence. The separation of satisfaction (with the hearing aid or with the service) from benefit (measured as a reduction in disability or in handicap) in the minds of clients, and the relative value of each of these concepts to them, must await further study.

CONCLUSIONS

Numerous measures of benefit from and satisfaction with rehabilitation can be shown to correlate with each other, thus lending validity to each of these measures. The measures that were most highly correlated with the others were the SHAPIE (a 25-item questionnaire addressing aid benefit), HAUQ3 (a 6-item questionnaire addressing aid benefit), ASS (a single-item questionnaire addressing satisfaction with the hearing aid), and the COSI (a five-, or fewer, item questionnaire in which the client suggests the items). Of these measures, the COSI is recommended as the most useful clinical measure of rehabilitation outcomes because it is not intrusive in the rehabilitation process and because its use has the potential to positively affect the rehabilitation process for individual clients.

Acknowledgment. We would like to thank Denis Byrne and Roger Lovegrove for helpful comments on the manuscript, Maria Massimino for assistance with administration of the experiment, and the numerous audiologists who participated in it.

REFERENCES

Byrne D, Dillon H. (1986). The National Acoustic Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid. *Ear Hear* 7:257–265.

Byrne D, Parkinson A, Newall P. (1990). Hearing aid gain and frequency response requirements for the severely/profoundly hearing impaired. *Ear Hear* 11:40–49.

Chermak G, Miller M. (1988). Shortcomings of a revised feasibility scale for predicting hearing aid use with older adults. *Br J Audiol* 22:187–194.

Cox RM, Alexander GC. (1995). The Abbreviated Profile of Hearing Aid Benefit. *Ear Hear* 16:176–186.

Cox RM, Gilmore C, Alexander GC. (1991). Comparison of two questionnaires for patient-assessed hearing aid benefit. *J Am Acad Audiol* 2:134–145.

Cox RM, Rivera IM. (1992). Predictability and reliability of hearing aid benefit measured using the PHAB. *J Am Acad Audiol* 3:242–254.

Demorest M, Walden B. (1984). Psychometric principles in the selection, interpretation, and evaluation of communication self-assessment inventories. *J Speech Hear Disord* 49:226–240.

Dillon H. (1994). Shortened Hearing Aid Performance Inventory for the Elderly (SHAPIE): a statistical approach. *Aust J Audiol* 16:37–48.

Dillon H, Koritschoner E, Battaglia J, Lovegrove R, Ginis J, Mavrias G, Carnie L, Ray P, Forsythe L, Towers E, Goulias H, Macaskill F. (1991a). Rehabilitation effectiveness I: assessing the needs of clients entering a national hearing rehabilitation program. *Aust J Audiol* 13:55–65.

Dillon H, Koritschoner E, Battaglia J, Lovegrove R, Ginis J, Mavrias G, Carnie L, Ray P, Forsythe L, Towers E, Goulias H, Macaskill F. (1991b). Rehabilitation effectiveness II: assessing the outcomes for clients of a national hearing rehabilitation program. *Aust J Audiol* 13:68–82.

Forster S, Tomlin A. (1988, May). *Hearing Aid Usage in Queensland.* Paper presented at the Audiological Society of Australia Conference, Perth.

Gatehouse S. (1993). Role of perceptual acclimatization in the selection of frequency responses for hearing aids. *J Am Acad Audiol* 4:296–306.

Gatehouse S. (1994). Components and determinants of hearing aid benefit. *Ear Hear* 15:30–49.

Kiresuk T, Sherman R. (1968). Goal attainment scaling: a general method of evaluating comprehensive mental health programmes. *Community Ment Health J* 4:443–453.

Lovegrove R, Battaglia J, Dillon H, Oong R. (1992). Report on outcomes assessment measures for new clients. *Aust Hear Serv Audiol Circ* 1992/7.

Malinoff R, Weinstein B. (1989a). Measurement of hearing aid benefit in the elderly. *Ear Hear* 10:354–356.

Malinoff R, Weinstein B. (1989b). Changes in self-assessment of hearing handicap over the first year of hearing aid use by older adults. *J Acad Rehabil Audiol* 22:54–60.

McKenna L. (1987). Goal planning in audiological rehabilitation. *Br J Audiol* 21:5–11.

Mulrow C, Aguilar C, Endicott J, Tuley M, Velez R, Charlip W, Rhodes M, Hill J, DeNino L. (1990). Quality-of-life changes and hearing impairment. A randomized trial. *Ann Intern Med* 113:188–194.

Mulrow CD, Tuley MR, Aguilar C. (1992). Sustained benefits of hearing aids. *J Speech Hear Res* 35:1402–1405.

Newman C, Weinstein B. (1988). The Hearing Handicap Inventory for the Elderly as a measure of hearing aid benefit. *Ear Hear* 9:81–85.

Schum D. (1992). Responses of elderly hearing aid users on the Hearing Aid Performance Inventory. *J Am Acad Audiol* 3:308–314.

Taylor K. (1993). Self-perceived and audiometric evaluations of hearing aid benefit in the elderly. *Ear Hear* 14:390–394.

Ventry IM, Weinstein BE. (1982). The Hearing Handicap Inventory for the Elderly: a new tool. *Ear Hear* 3:128–134.

Ventry IM, Weinstein BE. (1983). Identification of elderly people with hearing problems. *ASHA* July:37–42.

Walden B, Demorest M, Hepler E. (1984). Self-report approach to assessing benefit derived from amplification. *J Speech Hear Res* 27:49–56.

Weinstein BE, Spitzer JB, Ventry IM. (1986). Test–retest reliability of the Hearing Handicap Inventory for the Elderly. *Ear Hear* 7:295–299.

APPENDIX

CLIENT ORIENTED SCALE OF IMPROVEMENT

Name : ______________________ **Category.** **New** ______ **Return** ______
Audiologist : ______________________
Date : **1. Needs Established** ______________
2. Outcome Assessed ______________

Degree of Change

Final Ability (with hearing aid)
Person can hear
10% 25% 50% 75% 95%

SPECIFIC NEEDS

Indicate Order of Significance

Specific Needs	Worse	No Difference	Slightly Better	Better	Much Better	CATEGORY	Hardly Ever	Occasionally	Half the Time	Most of Time	Almost Always
☐											
☐											
☐											
☐											
☐											

NATIONAL ACOUSTIC LABORATORIES

Categories

1. Conversation with 1 or 2 in quiet
2. Conversation with 1 or 2 in noise
3. Conversation with group in quiet
4. Conversation with group in noise
5. Television/Radio @ normal volume
6. Familiar speaker on phone
7. Unfamiliar speaker on phone
8. Hearing phone ring from another room
9. Hear front door bell or knock
10. Hear traffic
11. Increased social contact
12. Feel embarrassed or stupid
13. Feeling left out
14. Feeling upset or angry
15. Church or meeting
16. Other