J Am Acad Audiol 10: 67–79 (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 Questionnaire (HAUQ)

Harvey Dillon*
Greg Birtles*
Roger Lovegrove*

Abstract

Self-report outcome measures were routinely collected from 4421 adult clients around Australia. The measures used were the Hearing Aid User's Questionnaire (HAUQ), which assessed hearing aid use, benefit, problems, and satisfaction, and the Client Oriented Scale of Improvement (COSI), which identified client needs, change in listening ability, and final listening ability in situations important to each client. Listening to television or radio and conversing with one or two others in a quiet place were the most frequently nominated needs. The benefit reported for noisy places was less than for quiet places, but very positive nonetheless. The normative data collected show a marked concentration of responses near the upper end of the scale for each of the outcome items except daily use. Consequently, correlations between the measures, although highly significant, were mostly less than 0.4. When the data were collapsed across subjects seen at each hearing center, benefit, satisfaction, usage, and problems with the hearing aids became much more strongly correlated with each other, with correlation coefficients up to 0.8. Benefit, satisfaction, usage, and the types of problems clients encountered with their hearing aids varied significantly from hearing center to hearing center. The most frequently reported problem was dissatisfaction with the quality of the subjects' own voices (i.e., occlusion effect), followed by feedback. The problems most closely related to usage, benefit, and satisfaction, however, were the presence of feedback, comfort of the earmold or earshell, and the quality of the users' own voices. These outcome measures appear to be most suitable for identifying needs, identifying individuals receiving markedly less than average benefit, and for finding small differences between outcomes for subgroups of the population.

Key Words: Benefit, hearing aids, outcomes assessment, rehabilitation, satisfaction, usage

Abbreviations: ANOVA = analysis of variance; APHAB = Abbreviated Profile of Hearing Aid Benefit; BTE = behind the ear; COSI = Client Oriented Scale of Improvement; GAS = Goal Attainment Scaling; HAUQ = Hearing Aid User's Questionnaire; HHIE = Hearing Handicap Inventory for the Elderly; ITE = in the ear; NAL = National Acoustic Laboratories; NAL-RP = NAL Revised Profound; SSPL = saturation sound pressure level

There is a growing recognition that the benefits of providing aural rehabilitation, including the provision of hearing aid(s), need to be quantified. The demand for this largely comes from providers of finance for rehabilitation, such as government and insurance companies. Perhaps the large differential in cost between the cheapest hearing aids and the most expensive is also exacerbating the need. It seems reasonable to ask that the increased cost of the more expensive devices be justified by demonstrating benefit additional to that obtained with less expensive devices. In this article, we

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

report on the experiences and data obtained using two outcome measures to monitor a national, large-scale rehabilitation program. The data can be considered to be normative data for these particular outcome measures as they were obtained under ordinary clinical conditions by many clinicians on clients who were not selected in any special manner. First, let us consider what properties an ideal outcome measure would have. Each of these properties is, in fact, applicable to the measurement of just about any quantity.

Efficiency

An outcome assessment measure needs to be efficient. The data must be able to be collected and analyzed in a time that is commensurate with the value of the information obtained. If not, then the measure will not be carried out, or else shortcuts will be taken that decrease the validity of the data.

Reliability

If the outcome measure is administered twice to the same person, then the difference in the results obtained must be small compared to the degree of benefit measured and also small compared to the spread of benefit observed among different individuals. If this is not the case, then there is little point in measuring the benefit for a particular person. It would be just as accurate to estimate that each person had a degree of benefit equal to the average benefit experienced by the population to which they belong. If the individual test–retest differences are large, then there may still be some value in measuring the benefit obtained by a population, so that at least the average benefit is known, even if we cannot say anything additional about the benefit received by an individual. There may still be good value in comparing the outcomes for one population to that for another population, such as people who use one type of hearing aid compared to those who use another, assuming that other factors such as hearing loss are held constant.

Validity

The outcome measure must assess what it purports to measure. The word "benefit" is often used (including in this paper), but it is a vague term. It may relate to a reduction in disability (e.g., Walden et al, 1984), meaning that the person is now more able to detect and understand sounds in particular circumstances. Alternatively, it may relate to a reduction in handicap (e.g., Ventry and Weinstein, 1982), meaning that the person's activities are now less restricted by their hearing impairment. Yet again, it may relate to a reduction in negative emotions that the person feels. Alternatively, the outcome measured may be something other than benefit. Two outcomes that have commonly been measured are usage and satisfaction. Usage can be justified on the basis that if a device is not proving to be beneficial it will not be worn. Satisfaction can be justified on the basis that it should be based on a comparison of perceived benefit versus benefit expected or required by the person. Any such discrepancy in benefit may be more important to the person than the actual benefit. Because benefit (each of the varieties), usage, and satisfaction are related but different, we would expect them to be partially correlated to each other. Dillon et al (1997) reported that the correlation between satisfaction and benefit was higher than that between usage and either of the other two. It is extremely difficult to establish the validity of a measure. Other than by looking at the items and seeing what they seem to be measuring (construct validity), validity can only be established by obtaining a sufficiently high correlation between the measure of interest and a second measure thought to be valid (predictive validity). Of course, this raises the problem of how the second measure was considered to be valid.

Interpretability

It is no use just being able to say to a person that the benefit they receive from a hearing aid is 3.5, even if that is a valid, reliable score that only took a few minutes to collect. The result is meaningless. To make it meaningful, it is necessary to ascribe words to the score, such as "this corresponds to a degree of benefit, averaged across situations, midway between 'helpful' and 'very helpful.'" The score becomes even more meaningful if one can say what proportion of the hearing-impaired population experiences this much, or more, benefit. Naturally, the same comments apply to satisfaction and usage.

Usefulness

There is no point in measuring anything if the measurement does not affect some subsequent action. Outcome measures can have two types of application. First, averaged across people in a defined population, scores can be used to justify the continuation, enlargement, or reduction of finance to support the rehabilitation, or may suggest to clinicians or researchers that rehabilitation using one device or method is superior to that using another device or method. Second, outcome measures obtained on an individual can have a direct and important result. For example, a score showing that a person received less benefit than 90% of comparable people should provoke a clinician to critically reexamine every aspect of the rehabilitation program provided thus far. Conversely, a high score would confirm to the clinician that it was indeed time for the program to end. Furthermore, outcome measures have the potential to be useful, even before the outcome scores have been obtained. There is rarely time in a clinical program to try all of the rehabilitation options (including different types of hearing aids, training in strategies for coping with hearing loss, and counselling in dealing with the emotional aspects of hearing loss) that might be useful for a client. Consequently, the clinician has to make informed choices about how best to spend the time available. Some outcome measures can provide detailed information about the needs of an individual client. In this case, the clinician can use this information to help decide what device or type of technology the client should get and how time should be spent with the client.

This paper reflects the experience of using outcome assessment measures within a large hearing rehabilitation program. Australian Hearing Services is a government-funded organization that provides aural rehabilitation services to age pensioners, veterans, and children. Since 1990, outcome data have been routinely collected. For some of this time, the measures used were the Hearing Aid User's Questionnaire (HAUQ) and Goal Attainment Scaling (GAS; Dillon et al, 1991). The GAS measure requires the clinician to negotiate an acceptable target level of hearing ability with the client prior to commencement of the rehabilitation program. Although many clinicians found this to be a useful process, many did not think it was sufficiently useful to justify the time required to do it, and some did not like the process at all. Consequently, the GAS was replaced with the Client Oriented Scale of Improvement (COSI) in 1995, and the two measures used since then (COSI and HAUQ) are believed to be sufficiently reliable, efficient, valid, interpretable, and useful to justify their continued use. At irregular intervals, a random sample of these data is collected from some of the 70 hearing centers across the country. This paper reports on the accumulated data obtained in two such samples. The validity and reliability of the COSI have been reported in Dillon et al (1997). The HAUQ was first reported by Forster and Tomlin (1988) but has evolved slightly with time and will be described in this paper.

The aim of this paper is to provide normative data for the COSI and HAUQ questionnaires under typical clinical conditions and to assess the usefulness of collecting such data, both for individual clinicians and for organizations as a whole.

METHOD

Subjects

Participants were not subjects in the usual experimental sense. The participants were 4421 adult clients who were receiving routine clinical services. Individuals on whom outcome data were to be forwarded to a central location for analysis were chosen randomly by computer and their computer files marked accordingly. For some smaller hearing centers, outcome data were to be returned for all clients seen within a defined period of time. Subjects had a mean age of 76 with a standard deviation of 9.1 years. Their three frequency average hearing loss (500, 1000, and 2000 Hz) in the better ear had the distribution shown in Figure 1. Fifty-six percent of the subjects were receiving a hearing aid for the first time. The type of hearing loss was not recorded but, based on previous surveys from the same population (e.g., Macrae and Dillon, 1996), would be predominantly sensorineural.

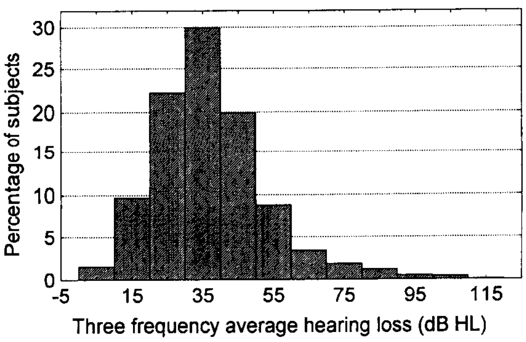

Figure 1 Distribution of three frequency average hearing losses in the better ear of the subjects.

Procedure

Because the data were obtained remotely from 46 different hearing centers around the country, the precise timing of data collection is not known. The first part of the COSI administration (identifying listening needs) was completed when the client was first assessed. The second part (quantifying the degree of benefit and residual difficulty in each situation) was completed at the end of the final regular appointment. Typically, the rehabilitation program is carried out in three appointments (primarily for assessment, fitting, and follow-up, respectively), each separated by about 4 weeks. We would therefore expect that the COSI would have been completed about 8 weeks after assessment and about 4 weeks after fitting, but a longer duration would have applied whenever additional appointments were needed to deal with problems arising from the hearing aid fitting or to provide additional counselling unrelated to hearing aids. The HAUQ data were collected at a short additional follow-up appointment 3 months after completion of the fitting program. Most commonly, this appointment was conducted by telephone, but some were conducted by having the client mail in the questionnaire, and some were conducted in person, at the discretion of the clinician. The number collected in each way is unknown.

After the subjects for whom outcome data were to be collected were randomly selected, arrangements were made to collect the HAUQ data. Because the first stage of COSI administration predated the random selection of which clients were to have their outcomes measured, COSI data were only available if the clinician had previously chosen to use COSI with the selected subjects. This was the case for 1770 of the 4421 subjects. The basis on which clinicians decided when to use COSI is largely unknown but is affected by clinicians' personal preferences and by time pressures. Approximately 200 clinicians provided the data reported in this paper.

The hearing aids fitted were predominantly (97%) digitally programmable single-channel hearing aids with output-controlled compression limiting to control SSPL. Some of the hearing aids were fitted with a more gradual form of compression activated in addition to the compression limiting. This more gradual compression had a 2:1 compression ratio and an adjustable compression threshold that typically was set to about 60 dB SPL. Some of the hearing aids had a second program made available to the wearer via a remote control or a switch on the hearing aid. This second program typically contained a low-frequency cut, relative to the main program. Sixty-seven percent of the hearing aids were in the ear (ITE) and 32.7 percent were behind the ear (BTE). The gain frequency response was selected using the National Acoustic Laboratories Revised Profound (NAL-RP) selection procedure (Byrne et al, 1990), which is identical to the NAL-R procedure (Byrne and Dillon, 1986) for people with mild and moderate losses. The NAL SSPL procedure (Dillon and Storey, 1998; Storey et al, 1998) was used to select maximum output.

COSI

The first step in performing a COSI assessment comprised identifying the specific listening situations in which each subject wished to be able to hear better. Up to five of these situations were recorded by the clinician. As part of the training package when COSI was introduced to the hearing centers, all clinicians had been instructed to question their clients sufficiently to obtain very specific listening situations. For example, "listening in noisy places" would not be considered sufficiently specific, whereas "talking to friends at the bar in the club on Saturdays" would be an extremely identifiable situation. The COSI form contains a list of 16 standardized listening situations (listed in Table 1). Each individual listening need is categorized into the most appropriate of these standard situations. This categorization does not affect the outcome score obtained for an individual, but it enables the listening needs of a population to be described and the degree of benefit obtained in a specific situation to be compared to that expected for the population listening in similar situations.

The second stage of COSI administration, which occurred after the clinician believed that all necessary services had been provided to the client, consisted of quantifying two things. The first was the degree of benefit experienced by the client in each of the previously identified listening situations. The "Degree of Change" section of the COSI form allows the following response choices: "worse," "no better," "slightly better," "better," and "much better." For the purposes of analysis, these were scaled as 1 to 5, respectively. The second thing quantified was the final ability of the subject to hear in each of the same listening situations. The allowable choices were "hardly ever," "occasionally," "half the time,"

Table 1 Number of Times Each of the 16 Standard COSI Needs Categories Was Mentioned at the Initial Appointment as a Percentage of the Number of Subjects

Category Number	*Category Description*	*Percentage of Subjects Mentioning*
1	Conversation with one or two in quiet	47.4
2	Conversation with one or two in noise	24.1
3	Conversation with group in quiet	31.9
4	Conversation with group in noise	23.5
5	Television/radio at normal volume	74.8
6	Familiar speaker on phone	10.9
7	Unfamiliar speaker on phone	5.0
8	Hear telephone ring from another room	9.3
9	Hear front doorbell or knock	1.8
10	Hear traffic	1.5
11	Increased social contact	1.8
12	Feel embarrassed or stupid	2.4
13	Feel left out	0.4
14	Feel upset or angry	0.6
15	Listening in church or meeting	19.8
16	Other	21.2

"most of the time," and "almost always." For those clients who prefer to answer in a more quantitative manner, each of these labels had an alternative numerical percentage. These were 10, 25, 50, 75, and 95 percent, respectively. The five response categories were again scaled from 1 to 5. In both cases, a higher number indicates a more successful outcome. The steps leading to the derivation of the COSI, and a form that facilitates its use, can be found in Dillon et al (1997).

HAUQ

The HAUQ is a multi-item questionnaire addressing several different issues. It is primarily designed to detect problems that may be affecting the person's ability to use and benefit from the hearing aid. The questions in the HAUQ are listed in the Appendix, and a copy can be obtained by writing to the authors. As can be seen, Questions 1 and 2 deal with usage of the hearing aid, with the categories in question 2 scaled from 1 to 6. Question 3 deals with benefits, with "not at all" scaled as a 1, "a little" scaled as a 2, and "a lot" scaled as a 3. Question 4 deals with problems, with "no" scaled as 2 and "yes" scaled as 1. Questions 5 to 7 deal with satisfaction, each scaled from 1 to 4. In all cases, the largest numbers correspond to the most favorable outcome. Question 8 seeks the clients' assessment of whether they have problems that require an additional appointment. Questions 9 to 11 are open-ended questions aimed at finding out what the clients liked and disliked about the service or devices they received.

RESULTS

COSI

Table 1 shows the number of times each of the 16 standard COSI need categories was mentioned at the initial appointment as a percentage of the number of subjects for whom COSI need data were available. The most commonly mentioned need was listening to television and/or radio, followed by conversation with one or two in quiet. Of the 16 standard need categories, the two least commonly mentioned were the two emotional needs: feeling left out and feeling upset or angry. The subjects mentioned an average of 2.8 needs per person. Only 11 percent of the subjects mentioned a single need, and only 5 percent of the subjects mentioned five needs (the maximum allowed for on the form).

Figure 2 shows the mean degree of change for each of the standard categories. A Kruskal Wallis analysis of variance (ANOVA) by ranks showed that the distribution of responses differed significantly across the categories ($p < .001$). (For this analysis, categories 9 to 14 were excluded because of the low number of responses within each.) Figure 3 shows the mean final listening ability for each of the categories. The distribution of responses again differed significantly across categories (Kruskal Wallis ANOVA, $p < .001$, with categories 9 to 14 excluded). Lis-

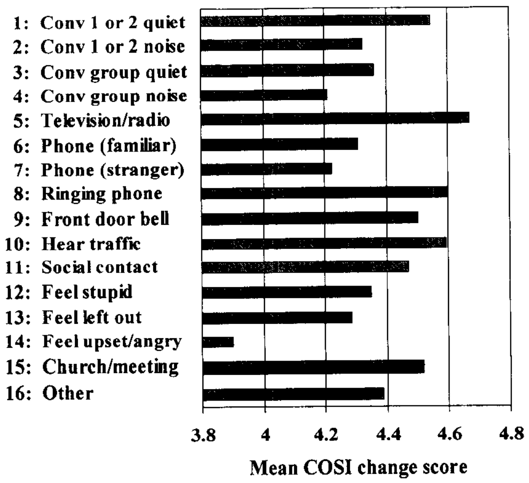

Figure 2 Mean COSI change scores for each of the 16 standard COSI listening situations.

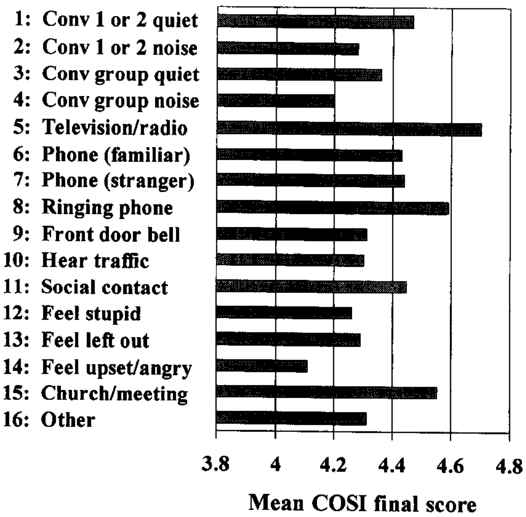

Figure 3 Mean COSI final listening ability scores for each of the 16 standard COSI listening situations.

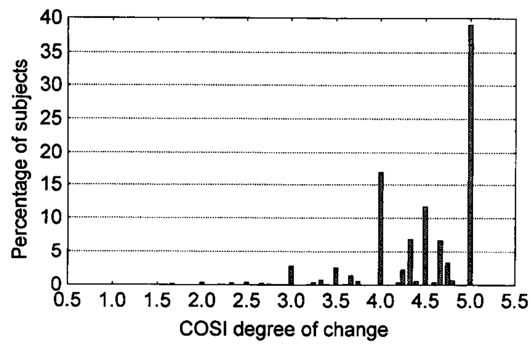

Figure 4 Percentage of subjects reporting each of the possible scores for the COSI change outcome measure.

tening to television and radio was the category with the biggest COSI change score and also the category with the highest final listening ability.

When the COSI change scores were averaged across needs within each subject, and then these mean scores were averaged across subjects, the overall mean change score was 4.47, which is midway between "better" and "much better." Similarly, the mean final listening ability was 4.45, which is midway between hearing "most of the time" and "all of the time." Table 2 shows the COSI change scores and final abilities exceeded for various proportions of the population. Eighty-nine percent of the subject population reported a mean improvement of "better" (a score of 4.0) or greater, and 40 percent reported a mean change of "much better" (a score of 5.0). This asymmetry of the distribution for the change score is clearly seen in Figure 4. Scores for the new and return clients have been combined as the ANOVA indicated that there was no significant difference between the scores obtained for the two groups.

Table 2 COSI Scores, Averaged across the Nominated Listening Situations for Each Person, That Are Equalled or Exceeded by the Percentage of the Population Indicated in the First Column

Population Percentile	*COSI Degree of Change*	*COSI Final Ability*	*HAUQ Benefit*
99	2.33	3.00	1.33
98	3.00	3.00	1.50
95	3.25	3.67	1.80
90	3.67	4.00	2.00
80	4.00	4.00	2.20
65	4.33	4.33	2.50
50	4.60	4.50	2.75
35	5.00	4.67	3.00
20	5.00	5.00	3.00

The final column shows corresponding data for HAUQ benefit (averaged across the items in question 3).

HAUQ

Of those fitted bilaterally, 80 percent reported wearing two hearing aids at the 3-month follow-up. The reported daily or weekly usage is shown in Figure 5. Close to 90 percent of the subjects reported wearing their hearing aids for more than 1 hour each day. The mean benefit reported for each of the six situations listed in HAUQ question 3 is shown in Figure 6. When averaged across the six situations, the mean benefit reported was 2.61, which is approx-

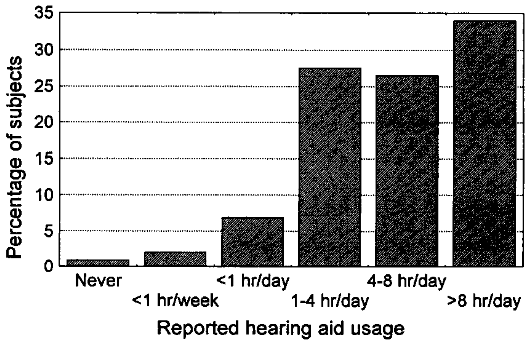

Figure 5 The percentage of subjects reporting different amounts of hearing aid usage per day or per week (based on HAUQ question 2).

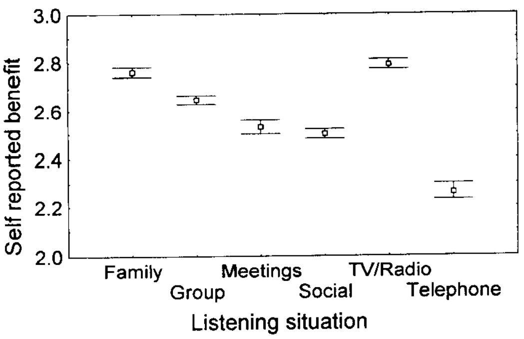

Figure 6 Mean reported benefit (based on HAUQ question 3) for each of the listening situations listed in question 3. The bars indicate ± two standard errors of the mean.

imately midway between "a little help" and "a lot of help." The benefit, averaged across the six listening situations, achieved by various proportions of the population is shown in the final column of Table 2. Only 16 percent reported a mean improvement of "a little" or less. As with the COSI data, the distribution is highly skewed toward the maximum possible benefit, with 41 percent reporting a mean benefit of 3.0 (the maximum possible).

Problems, defined as a "yes" response to any of the HAUQ question 4 items, were reported by 48 percent of the clients. Subjects reported an average number of 0.9 problems per person, with an interquartile range from 0 to 1 problems per person. Table 3 shows the percentage of clients who reported problems against each item. The most frequently occurring complaint (27%) was for the sound of their own voice, presumably indicating that occlusion was not adequately being controlled with venting or long canals on molds or shells. Occlusion problems were more frequent for ITE aids (32%) than for BTE aids (20%). The second most frequent complaint was for feedback whistling, where 19 percent of the subjects reported difficulties irrespective of whether they were fitted with ITE or BTE aids. Problems with both occlusion and feedback were reported by 8.5 percent of the subjects.

Table 3 Percentage of Subjects Reporting Difficulties with Their Hearing Aid (Based on HAUQ Question 4)

Difficulty	%
Positioning or removing aid	8.7
Adjusting controls	13.2
Whistling (feedback)	19.1
Shell or mold discomfort	12.4
Loud noises unbearable	6.1
Own voice quality	27.8

Although the HAUQ questionnaire was sent out well after the rehabilitation program was believed by the clinicians to have been satisfactorily concluded, 12 percent of clients indicated that they would appreciate receiving an additional appointment with the audiologist. Five of every six of these clients requesting an extra appointment had also reported one or more problems with their hearing aid in question 4. Satisfaction with the hearing aid(s) was high, and satisfaction with the way the clients had been treated was very high, as shown in Figure 7.

Relationships between Outcome Measures and Hearing Centers

Table 4 shows the significant Spearman's Rho correlations between most of the outcomes measures reported in the preceding two sections. The table also shows the correlation of each with three frequency average hearing loss in the better ear. Although there are many significant correlations ($p < .01$), all correlations are very low.

Scores were then averaged across all subjects seen at each of the hearing centers. This resulted in one score for each type of outcome measure for each hearing center. The different measures were then correlated with each other across hearing centers. Because some centers returned data for only a small number of subjects, centers were only included in each correlation if data on the measures being correlated had been returned for more than 50 subjects. For correlations involving the COSI data, 14 centers met this cri-

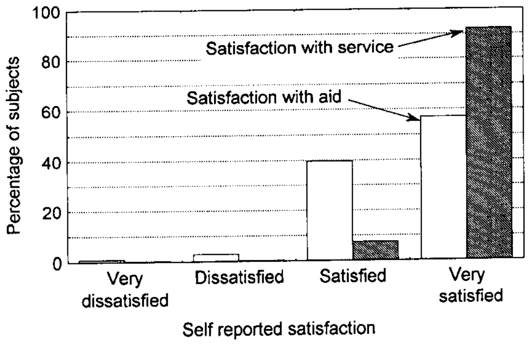

Figure 7 The percentage of subjects indicating different degrees of satisfaction with their hearing aid (based on HAUQ question 5) and with the service provided to them (based on HAUQ question 7).

Table 4 Spearman's Rho Correlation Coefficients between Three Frequency Average Hearing Loss (3FA) and Various Measures of Usage, Benefit, Problems, and Satisfaction

	3FA	*Usage (HAUQ 2)*	*Benefit (HAUQ 3)*	*Problems (HAUQ 4)*	*Aid Satisfaction (HAUQ 5)*	*Service Satisfaction (HAUQ 7)*	*COSI Change*	*COSI Final*
3FA	—	0.35	0.06			-0.05		
Usage	0.35	—	0.18	0.09	0.19	0.04	0.06	
Benefit	0.06	0.18	—	0.23	0.37	0.11	0.18	0.14
Problems		0.09	0.23	—	0.29	0.11		0.07
Aid satisfaction		0.19	0.37	0.29	—	0.26	0.16	0.15
Service satisfaction	-0.05	0.04	0.11	0.11	0.26	—		
COSI change		0.06	0.18		0.16		—	0.55
COSI final			0.14	0.07	0.15		0.55	—

Only correlations with $p < .01$ are shown.

terion, and for correlations involving only HAUQ data, 34 centers met the criterion. The resulting Pearson's correlation coefficients are shown in Table 5. The correlations are much more substantial, and all except those involving three frequency average hearing loss are in the expected direction: subjects from hearing centers where fewer problems are reported also report greater benefit, usage, and satisfaction. The negative correlations with three frequency average hearing loss are more puzzling. Each of these significant correlations was, however, totally caused by four hearing centers in one state that had poorer outcomes than average and more impaired subjects than average. There may, in fact, be no real correlation between three frequency average hearing loss and the other measures, at least over the range of mild and moderate hearing losses that dominate these data.

The high correlation between problems reported, satisfaction with the hearing aid, and reported benefit (HAUQ question 3) is particularly striking. The data are shown in Figure 8, with the rating of problems linearly transformed to show the average number of problems reported per client seen at each hearing center. To ascertain which of the problems mentioned in question 4 of HAUQ were most related to usage, satisfaction, and benefit, average ratings for each problem were correlated (across hearing centers) with ratings for hearing aid usage, benefit, and satisfaction. All correlations with benefit and satisfaction were significant, with correlation coefficients ranging from 0.41 to 0.78. Usage was significantly correlated with ease of control adjustment ($r = .47$), feedback ($r = .40$), and own voice quality ($r = .48$). Forward stepwise regression was used to find the problems that uniquely accounted for significant amounts of the variance in each of the dependent measures of usage, benefit, and satisfaction (HAUQ questions 2, 3, and 5). The results are shown in Table 6. Because all problems individually are correlated with the dependent measures, it would be unwise to dismiss any of the problems as unrelated to satisfaction, usage, or benefit, but from Table 6, mold/shell comfort, feedback, and own voice quality emerge as particularly important.

Table 5 Pearson's Correlation between Each of the Measures Listed in Table 4, Except the Data Have Been Collapsed across All Subjects Seen at Each Hearing Center

	3FA	*Usage*	*Benefit*	*Problems*	*Aid Satisfaction*	*Service Satisfaction*	*COSI Change*
3FA	—		-0.41	-0.42	-0.45		
Usage		—	0.43	0.42	0.48	0.37	
Benefit	-0.41	0.43	—	0.66	0.74	0.35	
Problems	-0.42	0.42	0.66	—	0.80	0.53	0.54
Aid satisfaction	-0.45	0.60	0.74	0.80	—	0.53	
Service satisfaction		0.47	0.34	0.53	0.53	—	
COSI change				0.54			—

Only correlations with $p < .05$ are shown.

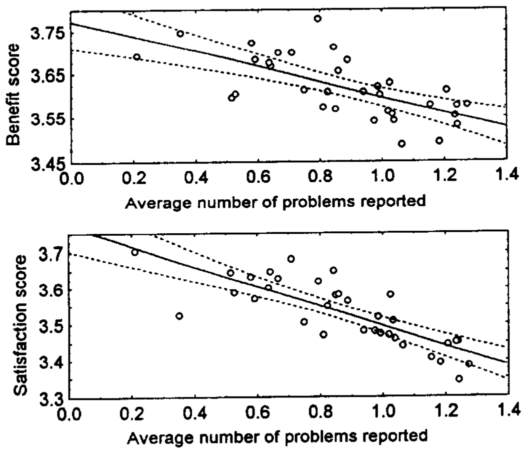

Figure 8 Benefit score (based on HAUQ question 3) and satisfaction score (based on HAUQ question 5) versus the average number of problems reported (based on HAUQ question 4) for each of 34 hearing centers that provided outcome data on more than 50 subjects. Solid line shows the regression and the dotted lines show the 95% confidence limits about the regression.

The correlation data suggest that subjects at different hearing centers report problems with different frequencies and different benefit and satisfaction outcomes. This was tested directly by performing a separate ANOVA for each type of outcome measure. In each case, the independent variable was "hearing center." Highly significant results ($p < .0001$) were reported for each of the dependent measures: usage, benefit, satisfaction, COSI change score, COSI final score, number of problems reported, incidence of the specific problems of control adjustment, feedback, mold/shell comfort, and own voice quality. In other words, there are significant differences in the reports coming from subjects from different hearing centers. As well as having statistical significance, many of them also had practical significance. The highest hearing center rating minus the lowest hearing center rating always exceeded a quarter of a rating point, which seems like a large change in an average score for scales with few points and a concentration of responses near the tops of the scales.

To test whether the pattern of problems complained about varied significantly from center to center, a two-way ANOVA was performed with hearing center as a between-groups factor and type of hearing aid problem as a repeated-measures factor. The analysis was restricted to those 34 hearing centers for whom more than 50 HAUQ returns were available. Both main effects were highly significant ($p < .000001$) and the interaction was also highly significant ($p = .0001$). For example, one hearing center received the third best rating (averaged across all subjects seen at that hearing center) for loudness discomfort but the worst rating for quality of own voice. There were numerous such examples of a hearing center apparently doing a much better than average job with one aspect of a hearing aid fitting but a much worse than average job of some other aspect.

DISCUSSION

Use of COSI

One of the interesting findings from this survey was the relative number of COSI and HAUQ data sheets returned by clinical staff. Administering the HAUQ was a mandatory request to the clinical staff, whereas the COSI form was requested to be returned if the first stage of the COSI had already been administered. Thus, it seems that clinical staff throughout the program are electing to perform COSI measurements on 40 percent of their clients. A reason commonly given by experienced staff for not performing a COSI evaluation is that the COSI process reflects what they normally do: find out what a person's hearing problems are and

Table 6 Results of Three Forward Stepwise Regressions with the Six Hearing Aid Problems Listed in HAUQ Question 4 as the Potential Independent Variables

Dependent Variable	*Variance Accounted for (R^2)*	*Significant Independent Variables*	*Significance Levels*
Usage	0.26	Own voice	.000001
		Feedback	.0001
Benefit	0.53	Feedback	.0005
		Mold/shell comfort	.0004
Satisfaction	0.67	Mold/shell comfort	.001

attempt to solve them. Some audiologists consider that filling in the COSI form adds nothing to this process (for them or for their client), so they elect never to complete the forms. Other audiologists consider that documenting the needs and the outcomes for each client adds to the effectiveness of their service, and they routinely administer COSI formally. A break-up of how many audiologists choose to use COSI always, sometimes, or never is unknown. Similarly, if there is any logical basis for choosing when to use COSI or when not to, that is also unknown. It seems unlikely that there would be any logical grounds for making this decision because the first part of COSI administration occurs very early in the clinical encounter, when the clinician knows very little about the client.

Compilation of data across people in a population is only slightly more difficult for COSI than for standard questionnaires. For standard questionnaires, particular items or subscales always relate to the same listening situation. The average benefit in particular situations can thus be estimated by averaging the responses across the items that relate to this situation. If the COSI is to be analyzed in this way, an additional step is necessary: each individual listening situation has to be categorized into the most appropriate standard situation, chosen from a list of commonly encountered situations. When this is done, COSI can also be used to describe the benefit for a population in a number of generalized listening situations (e.g., see Figs. 2 and 3).

Prior to collecting needs-based outcome data, we had hypothesized, based on numerous anecdotes, that a sizable minority of people would be better served by receiving an assistive listening device, or devices, rather than a regular hearing aid. Data of the type reported in this paper convince us that the proportion of people in this category is in fact very small. Although listening to television or radio was the most frequently stated need (see Table 1), only 1.8 percent of the clients reported a need for a device to help them with the television, telephone, or doorbell and did not nominate a conversation-oriented need (categories 1, 2, 3, 4, 15, or 16).

Hearing Aid Usage

Figure 5 showed that 90 percent of the clients reported using their aid(s) more than 1 hour per day. Similar studies carried out in Scandinavian countries report figures in the range of 64 to 75 percent (Warland and Tonning, 1988; Henrichsen et al, 1991; Vesterager et al, 1993). Although a self-reported usage of less than 1 hour per day should be sufficient grounds for the clinician to enquire whether there is anything wrong with the fitting, it should not be assumed that such hearing aid fittings are necessarily unsuccessful. Some subjects may have only needed help from a hearing aid for a small amount of time per day. For many of them, however, the low usage may have indicated that they were receiving little benefit from their fittings. Subjects who used the aid less than 1 hour per day (i.e., any of the first three response alternatives in HAUQ question two) reported an average satisfaction with the aid of 3.25, which was significantly ($p < .000001$) below the average satisfaction of 3.56 reported for the remaining subjects. Although statistically significant, this difference is only one-third of a scale point, so there are plenty of apparently satisfied people who only use their hearing aids for a small amount of time. Presumably, these people only needed help for a small amount of time. Although not shown, the average satisfaction grew linearly with the hearing aid usage category reported. On a note of caution, however, those people who stated that they never used their hearing aids reported an average satisfaction score of 2.97, which corresponds to "satisfied" on the 4-point scale. It is hard to see why many people who never use their hearing aids are "satisfied," so there may be a tendency for people to overstate their degree of satisfaction. The lack of a close relationship between usage and either satisfaction or benefit (despite the correlation between satisfaction and benefit) was noted by Dillon et al (1997).

A follow-up 3 months after the fitting revealed the high proportion of people who reported a problem at this time (48%). Table 3 shows that some problems are much more likely than others. Other researchers have also reported aid management problems when they surveyed clients some months after the fitting. Vesterager et al (1993) reported that 16 percent of the sampled clients experienced feedback, but Henrichsen et al (1988, 1991) reported higher figures of 22 and 30 percent. Warland et al (1988) found that 30 percent of their sample of ITC users complained about the quality of their own voice while wearing the aid. These authors also indicate high levels of satisfaction; 70 to 76 percent of clients were either satisfied or very satisfied with their hearing aids (Henrichsen et al, 1991; Vesterager et al, 1993). A high satisfaction rating cannot be taken to imply that all is well.

It cannot be assumed that a trouble-free exit from a rehabilitation program implies that the person will remain trouble free. It also cannot be assumed that people experiencing problems will seek to have them fixed. A massive 38 percent of the population reported one or more problems with their hearing aid in question 4 of the HAUQ but did not request an additional appointment, even when given the opportunity in HAUQ question 8. On the assumption that the problems reported could be solved for most of these people, it seems that a quick follow-up 3 months after fitting may represent a good investment of time.

Relationships between Outcome Measures

It will be apparent from the Procedure section of this paper that as an experiment, the procedural details were rather loosely controlled. The reason, of course, is that the data are not really from an experiment at all. The data analyzed in this paper were routinely collected by a large number of clinicians under all the time and other pressures of normal clinical life.

It is reassuring that under these conditions strong correlations between the measures can occur (when data are collapsed within a hearing center) and that different outcome measures (COSI change score and HAUQ question 3) can give similar conclusions about the relative benefit of hearing aids in different listening situations. The COSI data on how much benefit hearing aids give in different situations, when averaged across populations, agree well with the data from HAUQ question 3 (Fig. 2 and Fig. 3 vs Fig. 6). Both indicate that hearing aids work best for listening to television and radio and for conversations with smaller numbers of people, (especially in quiet places, see Fig. 2). Both indicate that the least benefit occurs for the telephone and for listening in situations where there are more than a few people. However, by no means should we conclude that hearing aids do not help in noisy places. For both measurement methods, considerable benefit was reported for larger and noisy gatherings of people; it is just that more benefit is reported for smaller and quieter gatherings.

Across subjects, correlations between the different measures are very low (see Table 4). One reason for this is not hard to see: for most of the measures, a large proportion of the subjects reported the maximum possible benefit allowed by the respective response scales. It is not possible to get large correlations when many subjects all give the same response. A different perspective on the relationship between the measures emerges when the same data are correlated after being collapsed across subjects seen at each hearing center. We now see correlations as high as 0.80. When the data are collapsed across a subgroup of subjects, there is no problem with individual responses being skewed toward the top of the response scale. If a particular part of the service is being carried out in a less than optimal manner for one subgroup of subjects, then fewer of those subjects will report the maximum benefit (or absence of problems) and the average statistic will reflect the quality of that part of the service.

Not surprisingly, it seems that the basics of hearing aid fitting have to be done well if clients are to report high satisfaction with the service, high benefit, high usage, and high satisfaction with the hearing aid (particularly the latter). Clients should be able to report no difficulties with positioning and removing the aid, adjusting the controls, feedback, mold or shell discomfort, loudness discomfort, and the quality of their own voices. Comfort with the mold/shell and freedom from feedback emerged as particularly important determinants of usage, benefit, and satisfaction (see Table 6).

Of course, achieving all of this is harder than simply saying it should be done. The ANOVA on each of the measures across centers indicated that subjects at some centers report significantly fewer problems, significantly more usage, significantly more benefit, significantly greater change in communication ability, and significantly greater satisfaction than those at other centers. The considerable variation in ratings between hearing centers suggests that there may be room for improvement. In some cases, particular hearing centers achieved considerably better than average ratings for some problems, but worse than average ratings for other problems. In other cases, particular hearing centers received above average ratings for most of the problems. Some received below average ratings for most of the problems. Identifying why particular centers are so good at dealing with problems common in hearing aids and transferring this expertise to other places would seem like a productive way to improve service quality, and hence client benefit and satisfaction.

It is possible that there is some other cause for the differences reported by subjects at different hearing centers, such as differences in the overall contentment of people in different regions or differences in expectations transmitted from the clinicians to the subjects. However, this seems much less likely than differences arising from how well they have been fitted, particularly as average ratings at particular centers are not generally high (or low) for all of the problems covered by question 4 of HAUQ. Indeed, the interaction between hearing center and incidence of problems reported was highly significant, showing that the types of hearing aid problems complained about varied significantly from center to center, and it is difficult to see any explanation for this except for the amount and quality of attention that staff in each center give to different aspects of the hearing aid fitting.

Usefulness of Outcome Measures

Let us conclude by asking the question posed earlier in this paper: are outcome measures like the COSI and HAUQ useful? We think the answer is yes, but depending on the use intended. First, if the outcome measure is intended to differentiate between a person who receives more than average benefit from one who receives only average benefit, then the answer is no. The poor correlations shown in Table 4 indicate that the measures are just not sensitive enough to accurately rank clients in terms of benefit. Similarly, we might expect the measures to be incapable of differentiating between two alternative rehabilitation options for an individual, unless one of the options is an extremely poor option. With the measures reported here, one fundamental reason for this insensitivity is that many subjects respond at or near the top of the response scale (see Figs. 4 and 7) and thus cannot be separated. The same conclusion can be reached by comparing the test–retest confidence interval of COSI scores (Dillon et al, 1997) to the difference between the median score and the score exceeded by only 10 percent of the population. The same conclusion is reached for the Abbreviated Profile of Hearing Aid Benefit (APHAB; Cox and Alexander, 1995). Similarly, Dillon et al (1997) demonstrated that when the Hearing Handicap Inventory for the Elderly (HHIE; Ventry and Weinstein, 1982) is used as an outcome measure, the 95 percent confidence interval around the benefit measured is comparable to the mean benefit obtained from rehabilitation. It seems that for the COSI and HAUQ in particular, but probably for outcome measures in general, it is difficult to reliably detect the difference between people receiving average benefit and people who receive more than average benefit.

Second, the preponderance of very positive responses with the COSI and HAUQ measures means that when a client gives a poor rating for benefit, usage, satisfaction, or final communication ability, the response should send a very clear warning to the audiologist that all is not well with the service provided to that individual. The normative data in Table 2 and Figure 7 should assist the audiologist in knowing just how unusual any apparently low rating is.

Third, if data are collated across subgroups of clients, such as all clients seen at a particular center (or by a particular audiologist!), they can be very sensitive indicators of service quality and outcomes for the client, as evidenced by the highly significant ANOVA result for problems reported and the highly significant correlations between the different measures (see Table 5 and Fig. 8). When used in this way, the COSI data can show what clients' needs are and how well those needs are being met, and the HAUQ data can identify which aspects of the hearing aid fitting are not being adequately dealt with (such as own voice quality, Table 3), and whether this varies across subgroups of subjects.

Fourth, if this collation of data is made available to clinicians, they can use it to alter the structure and content of appointments to give more emphasis to avoiding the difficulties most likely to occur.

In conclusion, outcome data of the type reported here appear to be useful for assessing the relative outcomes of different populations, for identifying the needs of individual clients, and for identifying individual clients whose needs are not being met or who are having problems with their hearing aids. These measures do not appear to be useful for differentiating between individuals who receive an average degree of benefit versus those who receive above average benefit.

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.

Cox RM, Alexander GC. (1995). The abbreviated profile of hearing aid benefit. *Ear Hear* 16:176–186.

Dillon H, James A, Ginis J. (1997). The Client Oriented Scale of Improvement (COSI) and its relationship to several other measures of benefit and satisfaction provided by hearing aids. *J Am Acad Audiol* 8:27–43.

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

Dillon H, Storey L. (1998). The National Acoustic Laboratories procedure for selecting the Saturated Sound Pressure Level of Hearing Aids: theoretical derivation. *Ear Hear* 19:255–266.

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

Henrichsen J, Noring E, Christensen B, Pedersen F, Parving A. (1988). In-the-ear hearing aids. The use and benefit in the elderly hearing-impaired. *Scand Audiol* 17:209–212.

Henrichsen J, Noring E, Lindeman L, Christensen B, Parving A. (1991). The use and benefit of in-the-ear hearing aids. A four year follow-up examination. *Scand Audiol* 20:55–59.

Macrae JH, Dillon H. (1996). Gain, frequency response and maximum output requirements for hearing aids. *J Rehabil Res Dev* 33:363–376.

Storey L, Dillon H, Yeend I, Wigney D. (1998). The National Acoustic Laboratories procedure for selecting the Saturated Sound Pressure Level of Hearing Aids: experimental validation. *Ear Hear* 19:267–279.

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

Vesterager V, Sibelle P, Boisen G. (1993). Effectiveness of rehabilitation with hearing instruments—results of three years of systematic follow up. In: *Recent Developments in Hearing Instrument Technology. Proceedings of the 15th Danavox Symposium.* Copenhagen: Stougaard Jensen, 465–468.

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

Warland A, Tonning F. (1988). An evaluation of the use and benefit of in-the-canal hearing aids. In: *Hearing Aid Fitting. Proceedings of the 13th Danavox Symposium.* Copenhagen: Stougaard Jensen, 347–358.

APPENDIX

Hearing Aid User's Questionnaire

1. Do you usually wear one or two hearing aids?
2. On average, how often do you wear your hearing aid?
   8 or more hours per day
   4 to 8 hours per day
   1 to 4 hours per day
   Occasionally (less than 1 hour per day but more than 1 hour per week)
   Seldom (less than 1 hour per week)
   Never wear the hearing aid.

   If you never wear the hearing aid, please tell us why.
3. How much has the hearing aid helped you with any of the following?
   (a) Family
   (b) Small-group conversation
   (c) Meetings (e.g., committees, church)
   (d) Social activities (e.g. shopping, bowls)
   (e) Television and/or radio
   (f) Telephone

   [The response choices for each of these are "a lot," "a little," "not at all," or "help not needed."]
4. Current difficulties with the hearing aid:
   (a) Do you have any difficulty positioning the hearing aid or removing it?
   (b) Do you have any difficulty adjusting the controls on the hearing aid?
   (c) Does the aid whistle when it is in your ear and set at a comfortable listening level?
   (d) Does the fit of the hearing aid or earmold in your ear cause you any discomfort?
   (e) Does the hearing aid make any sudden loud noises unbearably loud (not just annoying)?
   (f) Does the sound of your own voice sound hollow or like it is echoing?
5. How would you describe your satisfaction with your hearing aid?
   [The response choices are "very satisfied," "satisfied," "dissatisfied," or "very dissatisfied."]
6. How would you describe your satisfaction with your hearing aid repair service?
   [The response choices are "very satisfied," "satisfied," "dissatisfied," "very dissatisfied," or "have not needed any repairs."]
7. How would you describe your satisfaction with the way you have been treated by Australian Hearing Services Hearing Centers?
   [The response choices are "very satisfied," "satisfied," "dissatisfied," or "very dissatisfied."]
8. Do you feel you need an appointment with your audiologist soon?
9. The thing I liked best about the hearing aid or service was___.
10. The thing I liked least about the hearing aid or service was___.
11. If I were to make a change to the hearing aid or service, it would be ___.