

The Attitudes towards Loss of Hearing Questionnaire (ALHQ): A Comparison of Paper and Electronic Formats

Gabrielle Saunders*
Anna Forsline*
Peter Jacobs*

Abstract

The purpose of the study was to determine whether scores obtained on the ALHQ when completed in electronic format are the same as when completed in paper format. Four groups of 25 individuals participated. Each completed the ALHQ on two occasions in either its paper version, its electronic version, or both. The variance in ALHQ scores from the first to second administrations was compared across test groups. Data showed that the two forms of the questionnaire yielded equivalent scores but that completion in different modes on both occasions resulted in more variability in scores than completion in the same mode on both occasions. It is concluded that when comparing questionnaire data across administrations, the same response format should be used. Electronic completion took longer than paper completion, but it is concluded that the numerous advantages of electronic administration outweigh the disadvantages of additional completion time.

Key Words: Counseling, hearing aids, hearing aid outcome, questionnaires

Abbreviations: ALHQ = Attitudes towards Loss of Hearing Questionnaire; E = Electronic; P= Paper; PVAMC = Portland VA Medical Center

Sumario

El propósito de este estudio fue determinar si los puntajes obtenidos en el Cuestionario de Actitudes hacia la Pérdida de la Audición (ALHQ), cuando se completa en formato electrónico, son los mismos que cuando se completa en formato de papel. Participaron cuatro grupos de 25 individuos. Cada grupo completó el ALHQ en dos ocasiones, tanto la versión en papel, en la versión electrónica o en ambas. La variancia en los puntajes para el ALHQ, de la primera a la segunda administración, se comparó en todos los grupos de prueba. Los datos mostraron que las dos formas del cuestionario rindieron puntajes equivalentes, pero que al completarlos en diferentes modos en ambas ocasiones se produjo más variabilidad que cuando se completó en el mismo modo en ambas ocasiones. Se concluye que cuando se comparan los datos del cuestionario en las diferentes aplicaciones, debe utilizarse el mismo formato de respuesta. La ejecución en formato electrónico tomó más tiempo que la realización en papel, pero se concluye que las numerosas ventajas de la administración electrónica superan las desventajas de un tiempo de administración más prolongado.

Palabras Clave: Consejería, auxiliares auditivos, resultado del auxiliar auditivo, cuestionarios

Abreviaturas: ALHQ = Cuestionario de Actitudes Hacia la Pérdida de la Audición; E = Electrónico; P= Papel; PVAMC = Centro Médico VA de Portland

*National Center for Rehabilitative Auditory Research, Portland VA Medical Center, Portland, OR

Gabrielle H. Saunders, National Center for Rehabilitative Auditory Research, Portland VA Medical Center, 3710 SW US Veterans Hospital Road, Portland, OR 97213; Phone: 503-220-8262, ext. 56210; Fax: 503-721-1402; E-mail: Gabrielle.saunders@med.va.gov

The research reported here was supported by the Department of Veterans Affairs Health Administration, Rehabilitation Research and Development Grant # C2645R and the National Center for Rehabilitative Auditory Research, Portland VA Medical Center, Portland, Oregon.

Hearing aid outcomes measurement is fast becoming a necessary component of clinical service as audiologists are more often obliged to demonstrate the efficacy of their treatment, provide evidence for third-party payment, conduct cost-effectiveness analyses, and justify allocation of resources. The majority of hearing aid outcome measures are questionnaire based (e.g., Hearing Handicap Inventory for the Elderly, Ventry and Weinstein, 1982; Satisfaction with Amplification in Daily Life, Cox and Alexander, 1999; Client Oriented Scale of Improvement, Dillon et al, 1997). Questionnaires are particularly useful because they provide the users' perspective regarding outcome and supplement laboratory-based outcome measures, which we know do not fully explain real-life listening (Cox, 2003).

Questionnaires, however, tend to be underutilized because of the time-consuming and relatively cumbersome nature of the data collection process when compared to use of basic objective clinical measures. Such processes include the use of large quantities of paper, time for questionnaire completion, review of patient responses to ensure items have not been missed, hand scoring of responses and/or entry of responses into a database, and, finally, interpretation of the data. However, now that computers are highly accessible and are familiar to many individuals, electronic questionnaire administration is a viable alternative. The advantages of electronic administration over paper format are numerous. They include accurate labeling with participant name and completion date, absence of missed responses, immediate entry of responses into a database, elimination of data entry errors, the potential for larger font size making completion by elderly and visually impaired individuals easier, automated scoring, comparison of responses to previous questionnaire administrations or to population norms, immediate availability of reports that can include data from more than one source (e.g., questionnaire scores plotted in relation to a participant's age and hearing level) and savings on resources, both financial and environmental. Additionally, electronic questionnaires can be used to make sophisticated decisions that personalize a questionnaire by adapting it to specific

responses and skipping irrelevant questions.

There are, however, some potential disadvantages to electronic administration of questionnaires. For example, many individuals consider computers to be impersonal and some are unfamiliar with computers, while others might have negative attitudes toward them. These factors may each alter patient responses. Furthermore, while not necessarily being a disadvantage, electronic questionnaires usually present questionnaire items one at a time, while with paper questionnaires, all items are visible at the start. This, and the fact that patients cannot see their prior responses, might influence the pattern of answers.

For these reasons, the American Psychological Association developed guidelines for computer-based tests that state that before applying norms from conventional tests to computer-based tests, equivalency of the two must be established (American Psychological Association, 1986). This equivalency includes ensuring that means, dispersions, and distributions of scores for the two modes of presentation are approximately the same.

A number of studies have compared the equivalency of paper and electronic questionnaires and have evaluated patient preference for the response format. Studies in which correlations between scores on paper and electronic questionnaires have been published show Pearson *r*-values of between 0.54 and 0.90 (e.g., Pouwer et al, 1998; Bliven et al, 2001; Caro et al, 2001). Additionally, data have shown that two-thirds or more of subjects prefer the electronic format over the paper format (e.g., Velikova et al, 1999; Caro et al, 2001; Ryan et al, 2002; Cook et al, 2004). Even when subjects did not prefer the electronic form, studies have shown that between 76% and 99% still found it acceptable, or found the electronic version easy to use (e.g., Pouwer et al, 1998; Velikova et al, 1999; Bliven et al, 2001). Preference for the electronic form was not related to age, sex, familiarity with technology, educational level, visual impairment, reading level, or the presence of arthritis (Drummond et al, 1995; Pouwer et al, 1998; Velikova et al, 1999; Bliven et al, 2001). Completion time for the electronic version of the questionnaires was similar (Ryan et al, 2002) or slightly longer (Bliven et al, 2001; Caro et al, 2001) than for the paper versions. However, when manual

versus automated data entry and scoring are considered, presumably these differences become insignificant. Furthermore, because most electronic questionnaires do not allow missing responses, electronic questionnaires were more complete than paper versions (e.g., Caro et al, 2001; Hanscom et al, 2002; Ryan et al, 2002).

In our laboratory, we have developed the Attitudes towards Loss of Hearing Questionnaire (ALHQ; Saunders and Cienkowski, 1996; Saunders et al, 2005). It examines attitudes toward hearing loss and hearing aids on five scales: Denial of Hearing Loss, Negative Associations, Negative Coping Strategies, Manual Dexterity and Vision, and Hearing-Related Esteem. It was developed with two purposes in mind: first, as a tool to elucidate some of the underlying psychosocial issues that lead to the refusal to acquire or to use amplification and, second, as a counseling tool to address these issues. The questionnaire takes about ten minutes to complete and is available in two forms: one for non-hearing aid users and one for current users of hearing aids. They differ in the wording of just six questions. The internal consistency values of the scales is good: four of the five scales have Cronbach's α values greater than 0.80. The test-retest reliability of the scales is also good; r-values range from 0.88 to 0.65. The low interscale correlations show that each measures a different construct (Saunders et al, 2005).

Both a paper version and an electronic version of the ALHQ have been developed. The electronic version has all of the advantages over the paper version that were described above, but prior to recommending implementation of the electronic ALHQ, it is necessary to confirm equivalence of the two types of questionnaire format. This study was therefore undertaken in order to determine (a) whether the electronic version

yields the same results as the paper version and (b) whether participants are able to understand and complete the electronic version as easily as they can complete the paper version. The paper version is available in Saunders et al (2005), and both versions can be downloaded from the National Center for Rehabilitative Auditory Research Web site (<http://www.ncrar.research.va.gov>).

METHODS

Study Design

The purpose of the study was to determine whether scores obtained on the ALHQ differ based upon the method of completion, that is, paper versus electronic. There were four groups of 25 participants. Each participant completed the ALHQ on two occasions in either its paper version (P), electronic version (E), or both (see Table 1 for details). Participants in each group were matched as closely as possible on age, gender, and hearing aid user status. The variance in ALHQ scores from the first to second administrations was then compared across test groups.

Participants

Eighteen women and 82 men, aged between 46 years and 80 years, participated (mean: 65.6 years, standard deviation: 8.9 years). All participants were recruited via fliers posted around the Portland VA Medical Center (PVAMC) or via the PVAMC Audiology and Speech Pathology Clinic records and were tested at the National Center for Rehabilitative Auditory Research in Portland, Oregon. All participants signed an IRB-approved informed consent form prior to participating. Participants had symmetrical sensorineural hearing loss. Symmetrical hearing was defined as a difference of 15 dB HL or less between the left and right ear pure-tone air-conduction thresholds averaged at 0.5, 1, 2, and 4 kHz. Fifty-six individuals had never used a hearing aid; forty-four individuals were current hearing aid users. No participant underwent any significant hearing-related event, such as a hearing aid fitting or a visit to the audiology clinic, between administrations 1 and 2.

Table 1. ALHQ Completion by Subject Group

Subject group	Mode of completion	
	Occasion 1	Occasion 2
PP	Paper	Paper
PE	Paper	Electronic
EP	Electronic	Paper
EE	Electronic	Electronic

Test Measures

The following test measures were completed by all participants.

1. *Pure-tone audiometry, otoscopy, and tympanometry.* Air-conduction thresholds were measured bilaterally at octave frequencies between 0.25 and 8 kHz, along with interoctave frequencies of 1.5, 3.0, and 6.0 kHz. Otoscopy and tympanometry were conducted to check for cerumen and conductive pathology respectively. Any participants with conductive pathology were excluded from the study. For later analyses, a four-frequency pure-tone average (4F-PTA) was computed (4F-PTA = mean of left and right ear thresholds at 0.5, 1.0, 2.0, and 4.0 kHz).

2. *The Attitudes toward Loss of Hearing Questionnaire (ALHQ v2.1).* The ALHQ v2.1 is a 22-item questionnaire with five scales: Denial of Hearing Loss (six items), Negative Associations (four items), Negative Coping Strategies (eight items), Manual Dexterity and Vision (three items), and Hearing-Related Esteem (two items). Each questionnaire item consists of a single statement, such as "I try to avoid small talk because of my hearing difficulties." Participants state the extent to which they agree or disagree with the statement on a five-point scale ranging from "Strongly agree" to "Strongly disagree." Two forms of the ALHQ are available; one for nonusers of hearing aids and one for current users of hearing aids. The forms differ in the wording of six questions. For example, one item on the nonusers form is "I am pretty sure that I don't need hearing aids"; the equivalent item on the current users form is "I really don't think that I need my hearing aids." The ALHQ is scored such that a high score on any scale is indicative of a less favorable attitude, that is, denial rather than acceptance of hearing loss, negative associations with hearing aids, poor coping strategies, poor manual dexterity and/or visual acuity, and low hearing-related esteem. Thus, low scores are considered preferable to high scores in terms of probable hearing aid outcome.

All items in the paper version of the questionnaire are printed on a single side of 8.5 x 11 inch paper in Arial 12-point font. The questionnaire and scoring key are printed in Saunders et al (2005).

The electronic version of the ALHQ

consists of three modules: a Patient Information module, a Questionnaire module, and a Report module. All data are stored in a database for later analysis.

The Patient Information module is completed by the clinician prior to the patient's arrival. The clinician enters patient data including name, address, date of birth, and audiometric thresholds. Age and audiometric thresholds will be referenced when the program is used to generate counseling recommendations.

The Questionnaire module is completed independently by the patient. It consists of three simple instruction screens (one or two sentences per screen), followed by 23 question screens. The patient's response to the first question "Have you ever worn hearing aids before?" directs the computer program to select the non-hearing aid users or current hearing aid users form, as appropriate. Each screen shows a single question, along with the response scale. Participants use either the mouse or the keyboard to indicate their response. Following a response, the program automatically moves to the next question. Once all questions have been answered, participants see all of their responses on a single screen. At this point they can change any response by clicking on the item they wish to revise.

The Report module automatically generates a report once the questionnaire has been completed. At this time, the report consists of demographic information, raw ALHQ scores, percentiles scores, and a profile graph of scores in relation to normative data (see Appendix 1). In the future the report will also contain counseling recommendations.

Procedures

All participants underwent audiometric testing, and a case history was taken. They then completed the ALHQ. If completing the paper version, the form appropriate to their history of hearing aid use was provided. Following completion, the experimenter checked that participants had responded to all items. Any missed items were completed at this time. All testing and completion of the ALHQ was done without a family member present.

Participants completing the electronic version were instructed to sit at the computer

and follow the instructions on the screen. The program automatically selects the correct version based upon participant responses and will not progress to the next question until a response has been entered. During completion the experimenter monitored participants' progress to determine whether they experienced any difficulties using the computer program.

Between six days and fifteen days after the initial visit, participants returned to the laboratory to compete the ALHQ for the second time using the version appropriate to their group designation.

The time taken to complete both the paper version and the electronic version during both administrations of the questionnaire was noted for participants in the EP and EE groups.

RESULTS

Participant Matching

Table 2 shows the means and standard deviations of pure-tone thresholds averaged across ears of participants in each experimental group. A repeated measures analysis of variance (ANOVA), using test frequency as the repeated variable and experimental group as the between-subjects factor, showed that thresholds did not differ significantly across the experimental groups ($F[3,96] = 0.129, p = 0.943$). This confirms that participants in each group were well matched on audiometric thresholds.

The mean age of participants in each group was very similar: PP: 64.8 yr., SD = 9.1; PE: 65.7 yr., SD = 9.3; EP: 66.6 yr., SD = 8.8

and EE: 65.5 yr., SD = 8.6. A univariate ANOVA showed the ages of participants in the four experimental groups did not differ significantly ($F[3,96] = 0.19, p = 0.906$). Ideally the distribution of men and women would have been identical across groups. However, due to difficulties obtaining female participants, there were five women in three of the groups (PP, PE, and EE) and just three women in one group (EP).

Means and standard deviations of ALHQ scores from administration 1 are shown in Figure 1 for participants in each experimental group separately. A comparison of these baseline scores using a multivariate ANOVA revealed an overall group difference ($F[3,96] = 2.4, p = 0.003$). Tukey HSD post hoc comparisons showed this difference to be mediated by six between-group differences. These are shown on Figure 1 by horizontal lines located above the two groups that differ from each other; the p-value on the line shows the degree of significance of that difference. The group differences in ALHQ scores from administration 1 are primarily on the Hearing-Related Esteem scale.

The number of days elapsed between administrations for each participant group was compared. The mean time elapsed for the PP group was 10.2 days (SD = 3.0); for the PE group it was 9.7 days (SD = 3.1); for the EP group it was 9.1 days (SD = 2.9); and for the EE group the mean time elapsed was 9.5 days (SD = 3.0). Univariate ANOVA showed these differences were not significant ($F[3,96] = 0.6, p = 0.647$).

Overall, then, participants in each experimental group were well matched on pure-tone thresholds, age, baseline ALHQ scores, and time between ALHQ administrations.

Table 2. Mean Threshold at All Test Frequencies for Each Experimental Group

Subject group	Test frequency (kHz)								
	0.25	0.5	1.0	1.5	2.0	3.0	4.0	6.0	8.0
PP	21.9 (9.8)	23.3 (10.4)	26.2 (12.8)	32.7 (15.9)	42.6 (21.3)	52.5 (23.6)	55.2 (21.1)	57.0 (19.0)	60.0 (19.1)
PE	19.8 (10.0)	22.5 (9.8)	25.2 (10.2)	31.8 (18.5)	38.4 (21.4)	53.3 (20.5)	55.8 (18.8)	57.9 (19.1)	59.7 (23.6)
EE	17.8 (13.4)	23.7 (14.3)	26.1 (15.4)	29.3 (18.1)	38.8 (21.6)	45.8 (22.8)	55.8 (23.1)	60.2 (22.9)	60.8 (22.4)
EP	20.9 (9.5)	25.6 (11.6)	28.8 (15.3)	33.8 (16.3)	41.9 (20.5)	47.6 (20.1)	57.5 (20.8)	61.5 (23.6)	63.4 (21.4)

Note: Standard deviations are shown in parentheses.

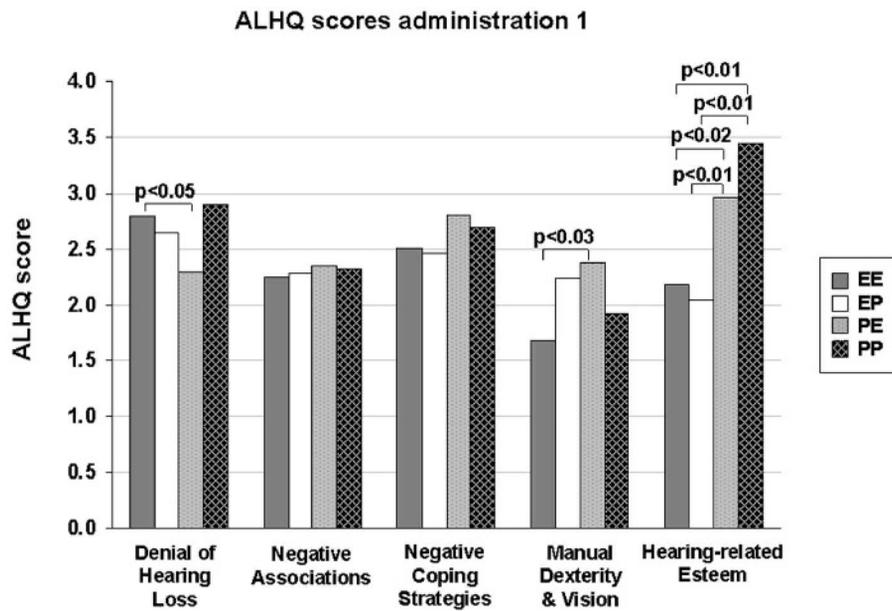


Figure 1. Bar graph showing group mean ALHQ scores at administration 1 along with horizontal bars showing scores that differ significantly at $p < 0.05$ or less. Dark shaded bars show EE group data; white bars show EP group data; light shaded bars show PE group data; and hatched bars show PP group data.

Effect of Demographic Variables upon ALHQ Scores

In order to determine whether the variables of age, sex, and hearing aid user status affected the data, three sets of analyses were conducted. For each of these analyses, the differences in ALHQ scores between administrations 1 and 2 were compared. Pearson correlations were used to examine the relationship between age and the difference in ALHQ scores between administrations 1 and 2. None of the correlations was significant at $p < 0.01$ (Denial of Hearing Loss: $r = 0.226$; Negative Association: $r = 0.01$; Negative Coping Strategies: $r = -0.116$; Manual Dexterity and Vision: $r = 0.003$; Hearing-Related Esteem: $r = 0.123$). Multivariate analyses of variance were then conducted to determine whether differences in ALHQ scores between administrations 1 and 2 differed by sex or hearing aid user status. Both ANOVAs showed nonsignificant group differences (Sex: $F[5,5] = 0.1$, $p = 0.992$; Hearing aid user status: $F[5,5] = 0.7$, $p = 0.577$). Since age, sex, and hearing aid user status were not significantly related to differences in ALHQ scores between administrations, the data from all subjects were combined for the between-participant group analyses.

Comparison of ALHQ Scores across Administrations

Pearson correlations were used to determine the strength of the relationship between ALHQ scores at administrations 1 and 2 for each experimental group separately. These correlations essentially examine test-retest reliability and were conducted to determine whether reliability differed across type of questionnaire administration (i.e., paper versus electronic). Confidence limits for these correlations were also computed. The results are shown in Table 3.

There is an indication that correlations differ across the participant groups. In general, the r -values for the PP and EE groups are higher than those of the PE and EP groups. In order to determine whether these differences were significant, a Fisher- z transform was conducted for each pair of correlations. If $z > 1.96$ for any comparison, then there is a significant difference between the correlations. The results of these computations are shown in Table 3 by the superscripted numbers 1 through 5. The correlations marked with the same superscript differ significantly from one another. There are five between-group differences: the correlations were significantly stronger among the PP group than the PE

group on the Denial of Hearing Loss scale ($z = 2.42$) and the Negative Associations scale ($z = 2.38$), and the test-retest correlations were significantly stronger among the EE group than the PE group on the Denial of Hearing Loss scale ($z = 2.96$), the Negative Associations scale ($z = 2.11$), and the Hearing-Related Esteem scale ($z = 2.00$). There were no other significant differences in correlations on any scales between the test groups. These data indicate that test-retest reliability decreases when different methods of questionnaire completion are used from administration 1 to 2. Note, however, there were no differences in the test-retest correlations of the EE and PP groups, or between the PE and EP groups.

To determine whether these differences in variability were related to a specific completion format (i.e., paper versus electronic) or to whether the format was the same or different between administrations, univariate ANOVAs were conducted comparing score at second administration across participant groups. Score at first administration was used as a covariate to account for differences in score at first administration. Results show nonsignificant between-group differences for all five scales (Denial of Hearing Loss: $F[3,95] = 0.8$, $p = 0.524$; Negative Associations: $F[3,96] = 2.1$, $p = 0.108$; Negative Coping Strategies: $F[3,96] = 1.1$, $p = 0.362$; Manual Dexterity and Vision: $F[3,95] = 1.6$, $p = 0.185$; Hearing-Related Esteem: $F[3,95] = 1.9$, $p = 0.136$). To illustrate these data further, Figures 2 and 3 were plotted. Figure 2 consists of histograms showing the distribution of differences in score between first and second administration

for each ALHQ scale and participant group separately. In each histogram, the values are normally distributed around the mean. Figure 3 shows group mean differences in score between first and second administration with error bars of ± 1 SD for each scale and participant group. Note that the mean group changes in score between administrations are small (within 0.25 points) and that Levene's Test of Equality of Error Variances shows that the variances were equivalent across all groups and ALHQ scales (Denial of Hearing Loss $F[3,96] = 0.5$, $p = 0.689$; Negative Association: $F[3,96] = 2.5$, $p = 0.063$; Negative Coping Strategies: $F[3,96] = 0.4$, $p = 0.766$; Manual Dexterity and Vision: $F[3,96] = 1.0$, $p = 0.379$; Hearing-Related Esteem: $F[3,96] = 0.6$, $p = 0.651$)

Completion Times

The time taken to complete the two forms of the questionnaire was compared using the data recorded from participants in the EP and EE groups at the second administration, that is, when participants in the EP group completed the ALHQ in its paper version, and participants in the EE group completed the electronic version. The mean time for the EP group was 4.9 min. (SD = 2.0 min.); the mean time for the EE group was 6.0 min. (SD = 2.8 min.). A univariate ANOVA showed this difference to be nonsignificant ($F[1, 47] = 2.2$, $p = 0.145$). However, when the time taken to complete the ALHQ at the first administration was used as a covariate (when both subject groups completed the ALHQ electronically), the difference between the groups became highly significant ($F[1,44] =$

Table 3. Correlations between ALHQ Scores at Administrations 1 and 2 in Bold and 95% Confidence Limits for Each Participant Group Separately

	Scale				
	Denial of Hearing Loss	Negative Associations	Negative Coping Strategies	Manual Dexterity and Vision	Hearing-Related Esteem
PP group (n = 25)	0.919 ¹ 0.831–0.962	0.916 ² 0.825–0.961	0.883 0.760–0.945	0.790 0.591–0.898	0.585 0.271–0.786
PE group (n = 25)	0.693 ^{1,3} 0.431–0.847	0.689 ^{2,4} 0.425–0.845	0.834 0.669–0.921	0.820 0.644–0.914	0.381 ⁵ 0.009–0.660
EP group (n = 25)	0.863 0.723–0.935	0.808 0.623–0.908	0.873 0.741–0.940	0.816 0.637–0.912	0.662 0.384–0.830
EE group (n = 25)	0.941 ³ 0.875–0.973	0.902 ⁴ 0.797–0.954	0.927 0.847–0.966	0.744 0.513–0.874	0.763 ⁵ 0.545–0.884

^{1, 2, 3, 4, 5} The difference between these pairs of correlations is significant at $p < 0.05$.

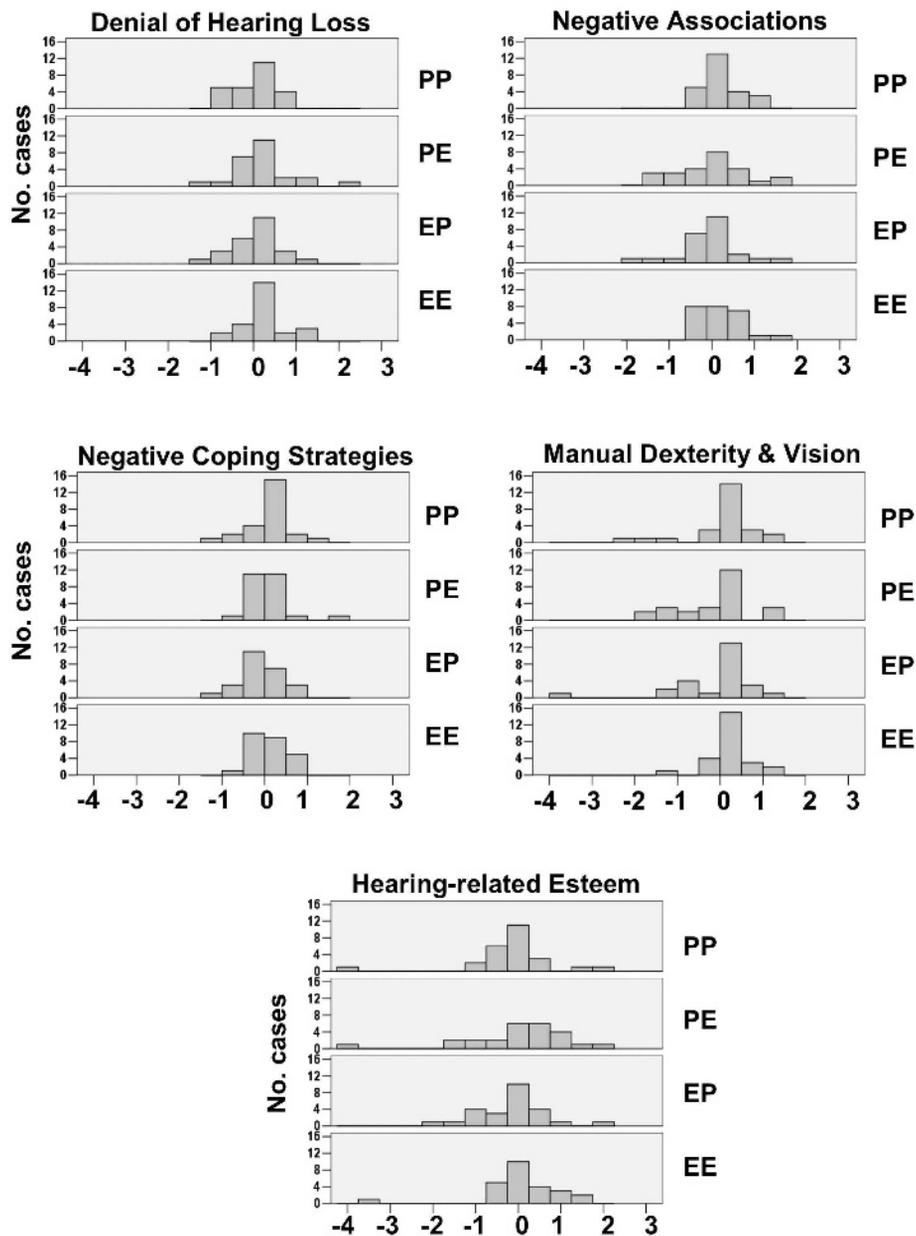


Figure 2. Histograms showing the distribution of differences in score between first and second administration for each ALHQ scale and participant group separately.

9.8, $p = 0.003$) such that the corrected time for electronic completion became longer relative to paper completion (Electronic: 6.2 min.; Paper: 4.6 min.).

Informal Observations

Informal observations made by the experimenter while subjects completed the questionnaire revealed that no participant

had difficulties comprehending or completing the paper version of the questionnaire. A few participants did require instruction from the experimenter on how to use the computer mouse, and some required clarification of where the “enter” key was on the keyboard. However, once these issues had been addressed, all participants were able to complete the electronic version easily and independently. No participants appeared to

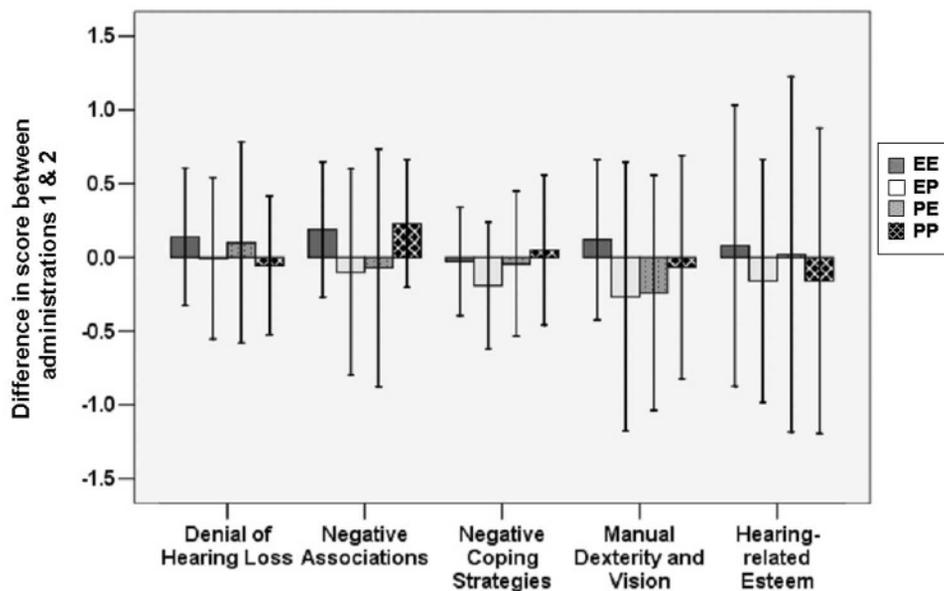


Figure 3. Bar graph of group mean differences in score between first and second administration for each scale and participant group along with error bars showing ± 1 SD. Dark shaded bars show EE group data; white bars show EP group data; light shaded bars show PE group data; and hatched bars show PP group data.

be anxious or stated that completion of either version made them anxious. Some individuals spontaneously reported that completing the questionnaire via a computer was more interesting than completing the paper version.

DISCUSSION

The primary purpose of this study was to determine whether completion of the ALHQ in electronic format yielded the same responses as completion of the ALHQ in paper format. In addition, completion time was recorded and informal observations were made in order to examine whether electronic completion raised additional issues such as difficulties handling a computer, additional time for completion, or anxiety with a nontraditional format. As discussed in the introduction, there are many advantages to having individuals complete questionnaires in an electronic form, but if by doing so different responses arise, then interpretation of scores would have to be modified, and new norms would have to be generated.

In order to examine the primary question regarding equivalence of scores with the two completion formats, comparisons of the correlations between scores at the first and second administrations for each participant group were made. These correlations ranged

from $r = 0.381$ to $r = 0.927$, which is similar to correlations shown in other investigations (e.g., Pouwer et al, 1998; Bliven et al, 2001; Caro et al, 2001). The rationale for comparing correlations between scores at the first and second administrations is that the correlations for individuals in the PP and EE groups who completed the ALHQ using the same format each time are a function of test-retest reliability only, whereas the correlations for individuals in the PE and EP groups who completed the ALHQ using a different format each time are a function of test-retest reliability combined with variability due to the change in response format. Thus, if the correlations for the PE and EP groups were significantly lower than those for the PP and EE groups, it would suggest that completion format impacted responses. The analyses revealed five between-group differences. In each instance the correlations that differed significantly were those between participants in the PE group, and participants in either the PP or EE group. This implies that completion format did impact response variability. Follow-up analyses were conducted to determine whether this variability was due to there being a difference in format between administrations (PE and EP versus PP and EE) or whether it was due to a specific format (E versus P). Since there were no between-group differences in scores at the second

administration, it is concluded that the differences in correlations were due to there being a difference in format between administrations rather than to the specific completion format. This should caution researchers and clinicians to ensure that the same response format is used any time data from questionnaires are to be compared.

In terms of missing data, the electronic ALHQ does not permit missing responses. Missing responses did occur with the paper version; however, they were not specifically tracked. The experimenter simply examined the questionnaire after completion and had the participant complete the missing items at that time. We are thus unable to provide information as to whether missing responses were question-specific or simply due to omission errors. Data regarding missing responses from other questionnaires completed in our laboratory are, however, available. We found that five of twenty-seven participants who completed a questionnaire known as the "Psychosocial Impact of Assistive Devices Scale" (PIADS; Day and Jutai, 1996) missed at least one question. This questionnaire is of particular relevance here because it has a very similar format to the paper version of the ALHQ. Further, across five different questionnaires completed (including the PIADS), nine of the twenty-seven participants missed at least one question. Thus, although these data are not questionnaire specific, they suggest that when completing questionnaires in paper format, participants do accidentally miss questions.

Some of the participants in this study required an explanation of how to use the computer mouse and keyboard. This might be because the population we used were relatively elderly: 27 of the 75 participants (36%) that completed the electronic ALHQ (subjects in groups PE, EP, and EE) were 70 years and older. Clearly, as the presence of computers proliferates, there will be fewer individuals unfamiliar with their use. Furthermore, once simple instruction lasting no more than two minutes had been provided, these individuals were able to use the computer program with ease.

The font size of the electronic version was considerably larger than the font in the paper version. This is advantageous for use with elderly visually impaired individuals. Moreover, because only one question at a time is shown in the electronic version, there

is no doubt that an individual is responding to that specific question, and not accidentally marking the wrong response line, as is possible with the paper version. On the other hand, as mentioned in the introduction, there is the possibility that the pattern of responses would differ because the participant cannot see their prior responses.

As in some previous studies (Bliven et al, 2001; Caro et al, 2001), completion time for the paper version was faster than for the electronic version by about one minute, or 20%. However, this time difference is compensated for by the time saved in the printing of questionnaires, manual data entry, and manual scoring. Furthermore, automated scoring decreases human scoring errors, and the availability of instant printed reports provides clinicians with a valuable resource not otherwise available.

CONCLUSIONS

This study shows that changing the response format of the ALHQ, and possibly other questionnaires, does affect the reliability of participant responses in that completion of questionnaires in two different modes resulted in more variability of responses between administrations than did completion in the same mode both times. However, this variability was not specific to the format of completion. Therefore, if responses from two or more administrations of a questionnaire are being compared, researchers and clinicians should ensure the same response format is used each time in order to minimize test-retest variability. Ideally, when designing an electronic version of a questionnaire, normative data should be collected to confirm that responses are the same as with the paper format. It is suggested that this extra work is worthwhile in light of the numerous advantages of electronic administration.

REFERENCES

- American Psychological Association. (1986) *Guidelines for Computer Based Tests and Interpretation*. Washington, DC: American Psychological Association.
- Bliven B, Kaufman S, Spertus J. (2001) Electronic collection of health-related quality of life data: validity, time, benefits and patient preference. *Qual Life Res* 10:15-22.

Caro J, Caro I, Caro J, Wouters F, Juniper E. (2001) Does electronic implementation of questionnaires used in asthma alter responses compared to paper implementation? *Qual Life Res* 10:683–691.

Cook A, Roberts D, Henderson M, Van Winkle L, Chastain D, Hammill-Ruth R. (2004) Electronic pain questionnaires: a randomized, crossover comparison with paper questionnaires for chronic pain assessment. *Pain* 110:310–317.

Cox R. (2003) Assessment of subjective outcome of hearing aid fitting: getting the client's point of view. *Int J Audiol* 42:S90–96.

Cox R, Alexander G. (1999) Measuring Satisfaction with Amplification in Daily Life: the SADL Scale. *Ear Hear* 20:306–320.

Day H, Jutai J. (1996) Measuring the psychosocial impact of assistive devices: the PIADS. *Can J Rehabil* 9:159–168.

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.

Drummond H, Ghosh S, Ferguson A, Brackenridge D, Tiplady B. (1995) Electronic quality of life questionnaires: a comparison of pen-based electronic questionnaires with conventional paper in a gastrointestinal study. *Qual Life Res* 4:21–26.

Hanscom B, Lurie J, Homa K, Weinstein J. (2002) Computerized questionnaires and the quality of survey data. *Spine* 27:1797–1801.

Pouwer F, Snoek F, van der Ploeg H, Heine R, Brand A. (1998) A comparison of the standard and the computerized versions of the Well-being Questionnaire (WBQ) and the Diabetes Treatment Satisfaction Questionnaire (DTSQ). *Qual Life Res* 7:33–38.

Ryan J, Corry J, Attewell R, Smithson M. (2002) A comparison of an electronic version of the SF-36 General Health Questionnaire to the standard paper version. *Qual Life Res* 11:19–26.

Saunders G, Cienkowski K. (1996) Refinement and psychometric evaluation of the Attitudes Towards Loss of Hearing Questionnaire. *Ear Hear* 17:505–519.

Saunders G, Cienkowski K, Forsline A, Fausti S. (2005) Normative data for the Attitudes towards Loss of Hearing Questionnaire. *J Am Acad Audiol* 16:637–652.

Velikova G, Wright E, Smith A, Cull A, Gould A, Forman D, Perren T, Stead M, Brown J, Shelby P. (1999) Automated collection of quality-of-life data: a comparison of paper and computer touch-screen questionnaires. *J Clin Oncol* 17:998–1007.

Ventry I, Weinstein B. (1982) The Hearing Handicap Inventory for the Elderly: a new tool. *Ear Hear* 3:40–46.

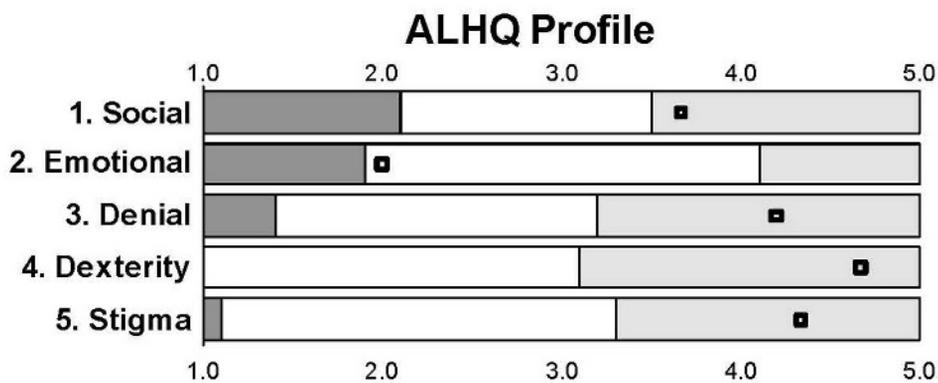
Appendix 1

ALHQ Report for Sample Subject

Clinic
 NCRAR
 3710 SW US Veterans Hospi
 Portland OR 97213

Patient	4-Frequency PTA
Sample Subject	Left: 50.0
DOB: 4/4/1944	Right: 51.3
PVAMC	
Portland	
OR 97207	
Non-Experienced Hearing Aid User	

ALHQ Summary Version 2.1			Completed: 2/16/2005		
<i>Subscale</i>	<i>Score</i>	<i>Percentile</i>			
Scale 1. Social impact of hearing loss	3.67	70 to 80			
Scale 2. Emotional impact of hearing loss	2.00	30 to 40			
Scale 3. Denial	4.20	80 to 90			
Scale 4. Manual dexterity	4.67	90 to 100			
Scale 5. Stigma associated with hearing aids	4.33	80 to 90			



Scores in the white region fall within 1 standard deviation of the mean.

Scores in the light shaded region are more than 1 standard deviation above the mean. Counseling with the associated module is recommended (See 'Recommendations', P2).

Scores in the dark shaded region are more than 1 standard deviation below the mean. Counseling may be recommended (See 'Recommendations', P2).