Original Article

International Journal of Audiology 2002; 41:30–35

Robyn M. Cox
Genevieve C. Alexander

The Department of Veterans Affairs Medical Center, Memphis, TN, and the University of Memphis, Memphis, TN, USA

The International Outcome Inventory for Hearing Aids (IOI-HA): psychometric properties of the English version

El Inventario Internacional de Resultados para Auxiliares Auditivos (IOI-HA): propiedades psicométricas de la versión en inglés

Key Words
Hearing aid
Hearing loss
Outcomes
Rehabilitation

Abstract

The International Outcome Inventory for Hearing Aids (IOI-HA) is a seven-item questionnaire designed to be generally applicable in evaluating the effectiveness of hearing aid treatments. The inventory was developed to facilitate cooperation among researchers and program evaluators in diverse settings. It is brief and general enough to be appended to other outcome measures that might be planned in a particular application, and will provide directly comparable data across otherwise incompatible projects. For this plan to be successful, it is essential to generate psychometrically equivalent translations in the languages in which hearing aid research and treatment assessments are performed. This article reports the psychometric properties of the inventory for the original English version. The items are reasonably internally consistent, providing adequate statistical support for summing the scores to generate a total outcome score. However, for maximum internal consistency, it would be desirable to generate two scores for the inventory.

Sumario

El Inventario Internacional de Resultados para Auxiliares Auditivos (IOI-HA) es un cuestionario de siete puntos diseñado para aplicarse de manera general en la evaluación de la efectividad de los tratamientos con auxiliares auditivos. El inventario se desarrolló para facilitar la cooperación entre los investigadores y quienes evalúan programas, en diversos ambientes. Por ser suficientemente breve y general puede ser agregado a otras formas particulares de medición de resultados que se apliquen y podrá proveer datos directamente comparables entre proyectos que de otra forma serían incompatibles. Para que este plan sea exitoso, es esencial generar traducciones psicométricamente equivalentes en las lenguas en las cuales se realiza la investigación sobre auxiliares auditivos y sobre el análisis de los tratamientos. Este artículo reporta las propiedades psicométricas del inventario en su versión original en inglés. Los puntos son razonable e internamente consistentes y proveen una base estadística adecuada para poder sumar los puntos y generar así un puntaje de resultados integral. No obstante, para lograr la máxima consistencia interna, es deseable que para el inventario se generen dos puntajes.

There is a growing awareness of the importance of the patient's point of view in determining the functional success of treatments in health-related fields. In the past, it was common for the success of an intervention to be judged by healthcare professionals, often based on laboratory or technical data. In the present, evaluations still take account of these types of data, but also factor in the judgement of the patient about the extent to which the treatment has alleviated the problems that he or she was experiencing in daily life. The upshot is that a treatment is unlikely to be seen as fully successful unless it can be shown to have resulted in an improvement in the patient's health-related quality of life. This paradigm shift is impacting on all branches of healthcare, including hearing healthcare. Thus, we are experiencing a burgeoning of interest in designing and using self-report instruments to document the patient's point of view in the evaluation of rehabilitative treatments using hearing aids (Bentler and Kramer, 2000).

The change in posture about the appropriate ways to judge hearing aid fitting outcomes is occurring simultaneously in numerous countries that deliver advanced healthcare. In a shrinking world where the same treatments are often used in different localities, both within and across countries, there are many valid reasons for scientists and practitioners to combine and compare data reflecting treatment outcomes. However, there are at least two formidable obstacles to combining and comparing outcome data that are based on patient self-report. First, there is no widely accepted standard self-report measure, so investigators and practitioners tend to adopt different measures. Even when different measures appear to address

Received:
July 12, 2001
Accepted:
September 27, 2001

Robyn M. Cox
Memphis Speech and Hearing Center,
807 Jefferson Ave,
Memphis, TN 38105, USA
E-mail: robyncox@latte.memphis.edu

similar issues and are written in the same language, they typically are dissimilar in subtle ways. As a result, data from various studies in the same language are often not directly comparable. Second, when investigations are pursued in different languages, the comparability of data is further compromised by unintentionally different nuances of meaning, even in questionnaire items that are intended to be the same.

Although it would facilitate combining and comparing data, it is not reasonable or desirable to propose a standard self-report inventory for evaluating hearing aid fitting outcomes. This would undoubtedly frustrate researchers, stifle innovation, and waste resources. Recognizing this problem, Cox et al (2000) proposed an alternative approach to achieving comparable data. They developed a short set of generally applicable items for self-assessment of hearing aid fitting outcomes, called the International Outcome Inventory for Hearing Aids (IOI-HA).

The IOI-HA is not intended to be used as a substitute outcome measure, but rather as a supplement. It is hoped that the IOI-HA could be appended to the battery of self-report measures that have been chosen for any research project. Because there are only seven items, this would absorb minimal additional resources. The payoff for including the IOI-HA in many different investigations would be the generation of a core of data that are directly comparable across diverse studies. This would allow the direct examination of issues such as the effectiveness of the same treatment when delivered in different healthcare contexts. It would also facilitate combining data from different studies to gain the additional power needed to test the significance of treatment effects or effect differences.

The original items for the IOI-HA were composed in English. The goals for the IOI-HA can be achieved only if there is a set of equivalent translations so that hearing-impaired individuals in different countries can complete the inventory in their native language. Further, it is highly desirable that there be only one translation for a given language, so that the psychometric properties of that version of the inventory can be clearly established, appropriate changes can be made as necessary, and there will not be confusion in the future when data obtained in a particular language are interpreted. Several members of the International Collegium of Rehabilitative Audiology (ICRA) undertook the task of generating a core of 'official' translations, and these are published elsewhere in this issue. To establish the comparability of the different translations, it is necessary to conduct a separate study of the psychometric properties of each translation. This article is the report of one such study for the English version. Reports of investigations of the Dutch translation, and of the English version administered in Wales, UK, are included in this issue.

Description of the IOI-HA

The inventory comprises seven items, each one targeting a different outcome domain. The domains are, in order: daily use, benefit, residual activity limitations, satisfaction, residual participation restrictions, impact on others, and quality of life. The wording and construction of items were chosen with the intention of minimizing literacy and cognitive demands. Each item has five response choices that are approximately semantically equidistant in English (Levine, 1981). The response choices always proceed from the worst outcome on the left to the best outcome on the right. The inventory is intended to be administered in paper and pencil mode. It is sufficiently self-explanatory that no formal instructions are needed. The items and response formatting are reproduced in the Appendix.

Method

The IOI-HA was mailed to 260 adults who had purchased hearing aids during 1999 or 2000 from one of two clinical service sites. The clinics were both private pay practices. An effort was made to contact every individual in the sampled group. No effort was made to segregate the subjects by type of instruments, or any other fitting variable. Subjects were encouraged to provide candid feedback and told that their responses would be anonymous. The response rate was 73% (190 responses). Usable responses were obtained from 172 subjects.

Subjects

The subjects were mostly elderly, with a mean age of 72 years (range: 26–98). Forty-two per cent were women. When asked to estimate their unaided hearing difficulties in daily life, 37% reported mild-to-moderate problems, 36% reported moderately severe problems, and 27% reported severe problems.

Sixteen per cent of the subjects had worn their current hearing aids for less than 3 months, 45% reported wearing their hearing aids for 3–12 months, and the remainder (39%) reported wearing their hearing aids for more than 1 year. Fifty-two per cent reported wearing their hearing aids for 8–16 h/day. The rest reported a variety of daily wear times, ranging from 'none' (2%) to 4–8 h/day (25%). Fifty-nine per cent of the subjects reported wearing two hearing aids for most of the time. Forty-five per cent reported that this was their first experience with hearing aids.

Results

Each item was scored from 1 to 5 for the responses from left (worst) to right (best), respectively. Thus, a higher score is indicative of a better outcome. The mean score for each item is shown in Figure 1. All the mean scores fall

between 3.5 and 4.1, somewhat above the middle of the scoring range. This seems to be indicative of a subject group that is relatively happy with their fitting outcomes, on the whole. Nevertheless, there is room for improvement in the scores, which is a desirable feature if the inventory is to be useful for discriminating among treatments.

The distribution of scores for each item is shown in Figure 2. This figure depicts the percentage of the time that each response level was used for each item. For two of the items, item 1 (use time) and item 4 (satisfaction), the most frequent response was the highest score. One item (item 5, residual participation restrictions) shows an asymptote at a score of 4. All other items show a maximum frequency for a score of 4. For all items, the two lowest responses (indicating poor outcomes) were each used for less than 10% of the time.

Overall, examination of Figures 1 and 2 indicates that this heterogeneous group of individuals with recently purchased hearing aids was generally favorably impressed, but not delighted, with the outcomes of the fittings.

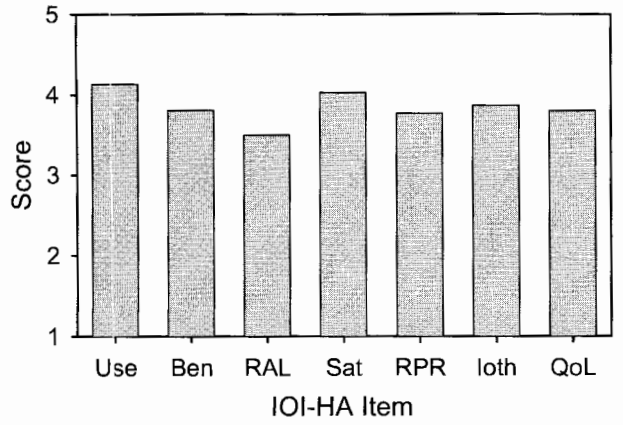

Figure 1. Mean score for each IOI-HA item. Ben, benefit; RAL, residual activity limitations; Sat, satisfaction; RPR, residual participation restrictions; Ioth, impact on others; QoL, quality of life.

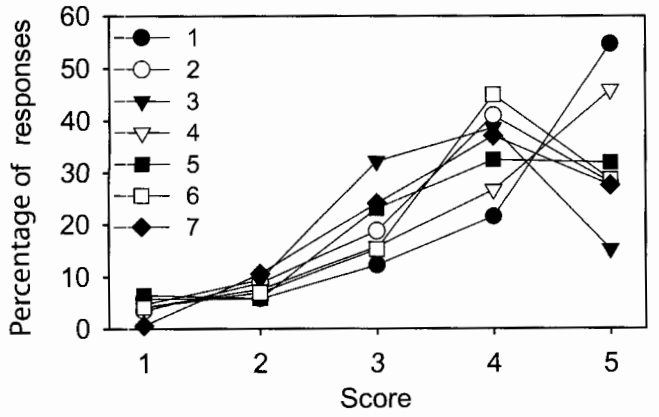

Figure 2. Distribution of responses for each IOI-HA item.

Although the goal of the inventory is to quantify the outcome of a hearing aid fitting from the patient's point of view, each of the items of the IOI-HA is intentionally devised to address a different domain of outcome data. It is of interest, therefore, to ascertain the extent to which the responses to the different items are related to each other. This was examined in three ways: inter-item correlations, factor analysis, and item–total statistics.

Table 1 depicts the inter-item correlations. They range from negligible (−0.04) to moderately strong (0.76). Further, the pattern of relationships is complex. Each item is significantly related to several other items, but none of them is related to all the other items. This result suggests that, despite the fact that all of the items tap some aspect of hearing aid fitting outcome, they might not actually be measuring the same underlying trait. This was further examined using a principal component analysis of the data. Parallel analysis was used to determine the number of extracted factors. Two factors were extracted, accounting for 46.7% and 22.1%, respectively, of the variance in scores. Table 2 gives the item loadings on the two factors after varimax rotation. Table 2 also

Table 1. Inter-item correlations for the English IOI-HA (n=168–171).

Item no.	*2 (Ben)*	*3 (RAL)*	*4 (Sat)*	*5 (RPR)*	*6 (Ioth)*	*7 (QoL)*
1 (Use)	0.39**	0.05	0.49**	−0.07	−0.04	0.45**
2 (Ben)		0.49**	0.65**	0.07	0.27**	0.68**
3 (RAL)			0.48**	0.28**	0.44**	0.43**
4 (Sat)				0.13	0.35**	0.76**
5 (RPR)					0.50**	0.08
6 (Ioth)						0.29**

Ben, benefit; RAL, residual activity limitations; Sat, satisfaction; RPR, residual participation restrictions; Ioth, impact on others; QoL, quality of life.
**Correlation is significant at the 0.01 level (2-tailed).

Table 2. Loadings of English IOI-HA items on each extracted factor after principal component analysis with varimax rotation (n=167).

Item no.	*F1*	*F2*
1 (Use)	0.73	
2 (Ben)	0.81	
3 (RAL)		0.62
4 (Sat)	0.86	
5 (RPR)		0.79
6 (Ioth)		0.82
7 (QoL)	0.86	
Coefficient α for factor	0.84	0.67

Loadings less than 0.5 are not shown. The last row gives the value of Cronbach's alpha for the combination of items in each extracted factor. Ben, benefit; RAL, residual activity limitations; Sat, satisfaction; RPR, residual participation restrictions; Ioth, impact on others; QoL, quality of life.

provides the value of Cronbach's alpha for each factor. Cronbach's alpha is a measure of the internal consistency of the items, or the extent to which they appear to measure the same attribute (Nunnally and Bernstein, 1994).

The factor loadings reveal a clear separation of the items into two factors. Factor 1 includes the items on use time, benefit, satisfaction, and quality of life. Factor 2 includes the items on residual activity limitations, residual participation restrictions, and impact on others. Factor 1 has been interpreted as encompassing introspection about the hearing aids ('me and my hearing aids'). Factor 2 is interpreted as reflecting the influence of the hearing aids on the individual's interactions with the outside world ('me and the rest of the world'). The results depicted in Tables 1 and 2 suggest that the IOI-HA is not a strictly unidimensional metric, because its seven items cluster into two clearly distinct areas.

Figure 3 depicts the distributions of scores for the two factors. Assuming no missing responses, the possible range of scores for Factor 1 is 4–20, and that for Factor 2 is 3–15. To facilitate interpretation, the score distribution for each factor was fitted with a best-fit third-order polynomial, and these are also shown in Figure 3. The polynomials explain 93% of the data for Factor 1 and 97% for Factor 2. Thus, they capture the trends in the data quite accurately. It can be seen that the distributions of scores for both factors span almost the entire range of possible scores, indicating that the inventory was effective in revealing a wide range of outcomes among subjects. The scores for the two factors have different distributions: for Factor 1, the most frequent outcome was near the top of the scale, with all other scores progressively less frequently seen: for Factor 2, the most frequent outcome was a score of about 12.

From a pragmatic point of view, the most straightforward method of dealing with responses to the IOI-HA items would be to sum the seven responses to produce a total outcome score, which would range from 7 to 35. Figure 4 illustrates the distribution of total outcome scores for the group of subjects in this study. Because the seven items do not reflect the status of a single underlying trait, there is room for debate about the legitimacy of combining the items together. One way to evaluate this is through item-total statistics. Table 3 gives two types of item–total statistics. The corrected item–total correlation shows the relationship between the score for the item and the total score for all the other items combined. Optimally, this relationship should be moderately high, and any item whose corrected item–total correlation is less than 0.20 should be eliminated from a combined score (Streiner and Norman, 1995). As shown in Table 3, none of the items has an unacceptably low item–total correlation, although item 5 is close to the cut-off.

The third column of Table 3 depicts the value of Cronbach's alpha for the scale if each item is deleted. The value of alpha for the seven-item scale is 0.78. if this value increases significantly when an item is removed, this indicates that the item is not very consistent with the other items, and suggests that it does not make a logical contribution to a score in which the items are combined. Once again, item 5 is seen to be anomalous in that alpha increased to 0.81 when item 5 was removed.

Combining the items into a total score allows us to evaluate the item characteristic curve (ICC) for each item. The ICC for an item shows the relationship between scores on the item and the total outcome score. This is useful in determining the extent to which each item helps to discriminate among individuals with different overall fitting outcomes. Figure 5 illustrates the ICCs for each of the seven items. The four items encompassed in factor 1 are in the top panel. The three items in factor 2 are in the lower panel. Only total scores that included at least five subjects were used in the ICCs. Because there were relatively few very low total scores, the ICCs could not be

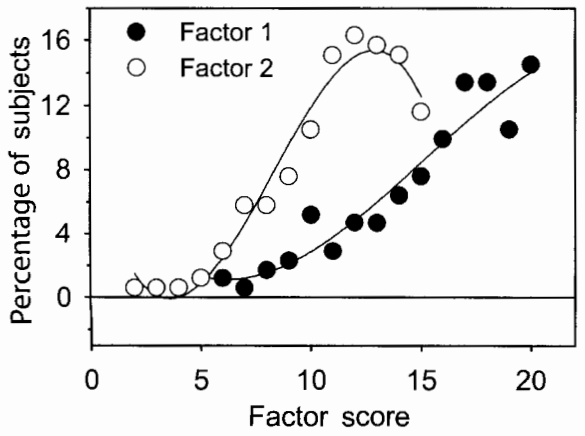

Figure 3. Distribution of IOI-HA outcome scores for Factor 1 and Factor 2.

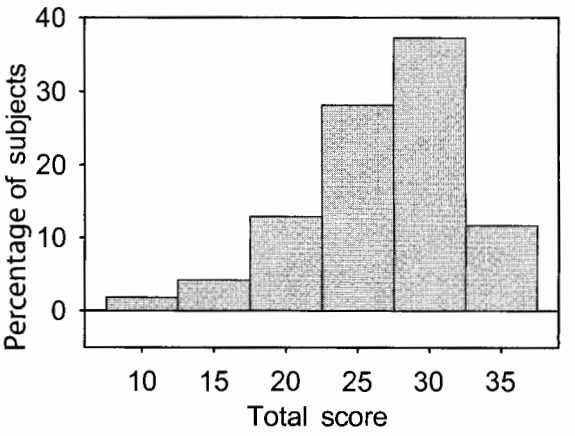

Figure 4. Distribution of total outcome scores for the IOI-HA.

Table 3. Item–total statistics for the English IOI-HA (n=167).

Item no.	*Corrected item-total correlation*	*Alpha if item deleted*
1 (Use)	0.30	0.79
2 (Ben)	0.65	0.72
3 (RAL)	0.54	0.74
4 (Sat)	0.75	0.70
5 (RPR)	0.22	0.81
6 (Ioth)	0.44	0.76
7 (QoL)	0.69	0.71

Ben, benefit; RAL, residual activity limitations; Sat, satisfaction; RPR, residual participation restrictions; Ioth, impact on others; QoL, quality of life.

generated for total outcome scores less than 20. To improve the interpretability of the figures, the data for each item were fitted with a second-order polynomial, and this is shown in Figure 5. For the items in the upper panel, the fitted curves each explain at least 95% of the variance in the data. In the lower panel, the fitted curves explain 93% (item 3), 92% (item 6), and 81% (item 5) of the variance in the data.

In an outcome measure such as the IOI-HA, where the goal is to separate individuals in terms of the degree of positive outcome, it is desirable for an ICC to rise monotonically to the maximum score without reaching an asymptote, as total score increases. This goal is clearly achieved for items 2, 3, 6 and 7, and mostly achieved for items 1 and 4. At least for total scores above 20, these six items do a good job of capturing different degrees of positive outcome. The ICC for item 5 is U-shaped, suggesting that high scores on this item can be associated with poor outcomes as well as with very good outcomes. Recall that item 5 queries residual participation restrictions (barriers to performing one's expected societal roles). This is a domain that is strongly influenced by the desired lifestyle of the patient. The implication of the ICC for item 5 is that patients whose lifestyle roles do not place much demand on hearing may get very little advantage from a hearing aid but still score highly on item 5. Further study is needed of this outcome domain to determine its potential heuristic value in outcome assessment.

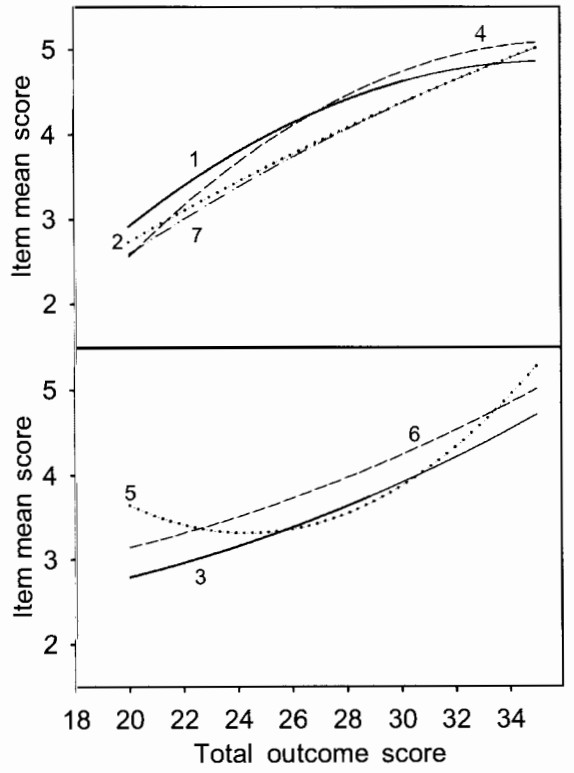

Figure 5. Item characteristic curves for each IOI-HA item.

Discussion

This investigation examined the psychometric characteristics of the English version of the IOI-HA by analyzing the responses given by a group of adults who purchased hearing aids in 1999 or 2000. This subject group was quite heterogeneous in terms of hearing problems and hearing aid styles, fittings, and technologies.

The generally high outcome scores shown in Figures 1, 2 and 4 indicate that the subjects in this study had relatively favorable attitudes towards their hearing aids. It is premature to conclude that these types of scores are ordinarily typical of recent purchasers of hearing aids. Recall that subjects were free to decline to participate in the study. It is possible that the self-selection process operated to reduce the proportion of respondents who were less pleased with their fitting outcomes. This potential problem is almost impossible to avoid in a retrospective study. An appropriately designed prospective investigation is needed to produce a distribution of outcome scores that might be indicative of all the hearing aid wearers in the sampled group.

If the data obtained in this investigation are typical of hearing aid wearers in general, the IOI-HA would appear to be well suited to detection of individuals who are negatively affected by their experience with amplification. Response distributions show that relatively few subjects selected the responses indicative of the poorest outcomes. Scores of 1 or 2 on an item or a total score of 20 or less are indicative of unusually pessimistic outcomes. Fewer than 15% of individuals reported outcome scores this low. On the other end of the scale, a total score of 33 or more identifies the individual as scoring in the top 10% of outcomes.

Given the results of this investigation, it is timely to consider how responses to the IOI-HA should be reported. Should the inventory be treated as a mini-profile in which each item is separately reported and, perhaps, compared with normative data? Should the item responses be summed to give an overall total score? Or should there

be two scores for the inventory, one for each factor? Arguments can be made for and against each one of these possibilities, and the best choice might depend on whether the inventory is being used for research, administrative or clinical purposes.

An overall total score is certainly the most parsimonious option and the simplest to interpret (although it is possible that important information might be lost in the summing process). This would seem to be a good choice if the IOI-HA is used administratively to document the outcomes of a service facility. A two-score index derived from the two factors identified (see Table 2) is clearly defensible, based on the item statistics. This might be the optimal choice when the IOI-HA is used as a research instrument. Finally, the approach that employs a mini-profile with norms could be quite useful for targeting areas in need of improvement for a particular individual if the inventory is used clinically to validate a fitting.

It is important to keep in mind that all the results reported here apply only to the English language version of the IOI-HA. It is highly desirable for all the translations of the inventory to produce similar psychometric data. Data for one translation are reported in this issue. Additional studies are needed to determine the psychometric characteristics of the other translations.

Appendix–English IOI-HA

1. Think about how much you used your present hearing aid(s) over the past two weeks. On an average day, how many hours did you use the hearing aid(s)?

None	Less than 1 h/day	1–4 h/day	4–8 h/day	More than 8 h/day
☐	☐	☐	☐	☐

2. Think about the situation where you most wanted to hear better, before you got your present hearing aid(s). Over the past two weeks, how much has the hearing aid helped in that situation?

Helped not at all	Helped slightly	Helped moderately	Helped quite a lot	Helped very much
☐	☐	☐	☐	☐

3. Think again about the situation where you most wanted to hear better. When you use your present hearing aid(s), how much difficulty do you STILL have in that situation?

Very much difficulty	Quite a lot of difficulty	Moderate difficulty	Slight difficulty	No difficulty
☐	☐	☐	☐	☐

4. Considering everything, do you think your present hearing aid(s) is worth the trouble?

Not at all worth it	Slightly worth it	Moderately worth it	Quite a lot worth it	Very much worth it
☐	☐	☐	☐	☐

5. Over the past two weeks, with your present hearing aid(s), how much have your hearing difficulties affected the things you can do?

Affected very much	Affected quite a lot	Affected moderately	Affected slightly	Affected not at all
☐	☐	☐	☐	☐

6. Over the past two weeks, with your present hearing aid(s), how much do you think other people were bothered by your hearing difficulties?

Bothered very much	Bothered quite a lot	Bothered moderately	Bothered slightly	Bothered not at all
☐	☐	☐	☐	☐

7. Considering everything, how much has your present hearing aid(s) changed your enjoyment of life?

Worse	No change	Slightly better	Quite a lot better	Very much better
☐	☐	☐	☐	☐

Acknowledgements

This article is based upon work supported by the Office of Research and Development, Rehabilitation R&D Service, US Department of Veterans Affairs (VA). Vickie Wilkinson and John Tecca assisted with subject recruitment.

References

Bentler RA, Kramer SE. (2000) Guidelines for choosing a self-report outcome measure. Ear Hear 21:37S–49S.

Cox RM, Hyde M, Gatehouse S, et al. (2000) Optimal outcome measures, research priorities, and international cooperation. Ear Hear 21:106S–15S.

Levine N. (1981) The development of an annoyance scale for community noise assessment. J Sound Vibration 74:265–79.

Nunnally JC, Bernstein IH. (1994) Psychometric theory, 3rd edn. New York: McGraw-Hill.

Streiner DL, Norman GR. (1995) Health measurement scales, 2nd edn. New York: Oxford University Press.