The purpose of this study was to examine the readability of 4 self-report measures commonly used in the field of rehabilitative audiology. The text from 4 self-report assessment tools used to assess hearing difficulty for adults and older adults (Hearing Handicap Inventory for the Elderly, Hearing Aid Performance Inventory, Hearing Handicap Inventory for Adults, and Abbreviated Profile of Hearing Aid Benefit) were analyzed using the Flesch Reading Ease Level (FRE), FOG, and FORCAST formulas. The reading levels on these self-report assessment tools exceeded the recommended reading level for reading materials in the health professions.

In the field of audiology, self-report measures are used to determine the need for hearing rehabilitative services and to measure the effectiveness of rehabilitative services given (Bentler & Kramer, 2000). The use of self-report tools in audiology is not a new concept. It has long been established that self-report tools provide helpful subjective in addition to objective information gathered during the audiologic assessment (Brainerd & Frankel, 1985; Demorest & Walden, 1984). Erdman (1994) states that self-report tools assess the typical performance of a client, while objective measurements such as audiologic assessments represent the maximum potential of a client. The focus of self-report assessments for this
Discussion is on those that assess the perception of hearing problems and hearing aid benefit by adults and older adults. According to the American Speech-Language-Hearing Association (ASHA; 1998), audiometric data alone are not sufficient for determining candidacy for rehabilitation for adults. The Ad Hoc Committee on Hearing Aid Selection and Fitting recommends a client-centered approach for adults that includes the assessment of the effects of hearing impairment at the personal activity and the social role level. That is, the use of self-report assessment tools is essential in establishing candidacy for amplification and rehabilitation. In addition, these self-report assessment tools are essential in the validation of the intervention provided to the adult client. Validation should also be conducted from a client-centered perspective with the use of self-report assessment tools.

The World Health Organization’s (WHO) International Classification of Functioning, Disability, and Health (ICF) framework (2001) provides a hierarchical, complex model to describe functioning and disability relative to any health condition. The ICF model has been used to describe the experiences of adults with communication disorders and their significant others (e.g., Brown & Hasselkus, 2008; Hickson & Scarinci, 2007; Hickson & Worrall, 2001; Scarinci, Worrall, & Hickson, 2009). The ICF has two domains and each domain has two components. Body functions and structures, and activities and participation, are the two components in one domain. An activity refers to the completion of a task or action. The negative expression of activity is an activity limitation which occurs when an individual has problems in the completion of the task/action. An example of an activity limitation in audiology would be difficulty understanding speech in the presence of background noise. Participation refers to involvement in a life situation. The negative expression of participation is a participation restriction which occurs when an individual has problems participating in life events. An example of a participation restriction in audiology would be not participating in group conversations because of the difficulty experienced conversing in background noise.

Environmental and personal factors are the two components of the other domain. Environmental factors include physical and social elements, as well as attitudes of those in the person’s environment. Personal factors refer to various aspects of the person’s life and experiences. These may include the individual’s age, gender, motivation, and personality. The ICF model emphasizes the complex, interactive nature of these components. That is, because of the differences in environmental and personal factors, there is not a one-to-one relationship between activity limitations and participation restrictions. It is therefore important to assess each of these four components. When working with adults, self-report assessments can be helpful tools in assessing these components of functioning.

Provided that most self-report assessment tools are designed to be administered in paper-and-pencil format, it is important to consider the ability of clients to accomplish this task. It is estimated that one-quarter of adults in the United States...
cannot read written materials that only require basic reading proficiency skills (Weiss & Coyne, 1997). According to the National Adult Literacy Study (National Center for Education Statistics, 1993), the average American adult reads at about the seventh grade level. Researchers in health literacy recommend that materials for the public be written at the fifth- or sixth-grade reading level (Doak, Doak, & Root, 1996; Weiss & Coyne, 1997). Not only would this practice be helpful for those with lower reading skills, but lower reading levels are preferred by readers of all levels (Weiss & Coyne, 1997). However, many written materials in the medical field are written at a tenth grade reading level or higher (Weiss & Coyne, 1997). This incongruity between reading ability, reading preferences of the general public, and reading level of printed materials can affect results of self-reported outcome questionnaires (Weiss & Coyne, 1997).

ASHA emphasizes the vision of making effective communication a human right, accessible and achievable for all (ASHA, 2010). ASHA further encourages members to learn more about health literacy because understanding health information is vital to one’s well-being. One way to address the issue of health literacy is to examine the readability of the materials used in a clinical setting. In the Healthy People 2010 document (2011), readability is listed as a key attribute of effective health communication. However, a discussion on readability of self-report assessment materials in audiology must be preceded by a discussion of the concept of readability.

The concept of readability is complex, but may refer to the ease with which the text is understood due to writing style (Klare, 1963), clarity (Hargis et al., 1998), reader characteristics (McLaughlin, 1968), or the combination of several factors which culminate in the success of readers to understand a text, read it with optimal speed, and find it interesting (Dale & Chall, 1948). The reading level of a text depends, in part, on the purpose and use of that text (DuBay, 2004). For example, a text used for independent, recreational use will have a higher reading level than a text used for learning purposes (DuBay, 2004). Self-report assessment tools align more closely with texts that are used for learning purposes. That is, these tools are used to elicit responses from a client who may or may not be familiar with the subject matter. In addition, clients have noted that they gain insights about themselves and their hearing loss simply by completing the questionnaires.

The readability of texts has been examined in the communication disorders field (Aleligay, Worrall, & Rose, 2008; Atcherson, Zraick, & Brasseux 2011; Greywoode, Bluman, Spiegel, & Boon, 2009; Kelly, 1996; Kelly & Kahn, 1991; Nair & Cienkowski, 2010; Zraick & Atcherson, in press). These studies concluded that most documents they examined exceeded the recommended fifth- to sixth-grade reading level. To date, however, no studies have been published which examine the readability of self-report hearing disability measures. The purpose of this study is to examine the readability of four commonly used self-report measures used to assess hearing disability in adults and older adults.
Materials

A review of the literature reveals that there are many self-report assessment tools used in audiology, and in rehabilitative audiology in particular. During the time of this study, the most-commonly cited self-report assessment tools were examined. The four most-common self-report assessment tools were: (a) Hearing Handicap Inventory for the Elderly (HHIE; Ventry & Weinstein, 1982), (b) Hearing Aid Performance Inventory (HAPI; Walden, Demorest, & Hepler, 1984), (c) Hearing Handicap Inventory for Adults (HHIA; Newman, Weinstein, Jacobson, & Hug, 1991), and (d) Abbreviated Profile of Hearing Aid Benefit (APHAB; Cox & Alexander, 1995).

The HHIE was designed to be administered to adults over the age of 65 years. The HHIE: (a) contains 25 items, (b) produces two subscales: one assesses the emotional consequences of hearing impairment and the other assesses social and situational effects, and (c) asks respondents whether a statement applies to them on a 3-point scale. The HAPI was designed to measure success with amplification. The HAPI: (a) contains 64 items; (b) contains four groupings of items: noisy situations, quiet situations with a talker nearby, reduced speech cues, and nonspeech stimuli; and (c) asks the respondent to rate the benefit received from hearing aids on a 4-point scale. The HHIA was modified from the HHIE to be used with adults under the age of 65 years. The HHIA: (a) contains 25 items, (b) produces two subscales: one assesses the emotional consequences of hearing impairment and the other assesses social and situational effects, and (c) asks respondents whether a statement applies to them on a 3-point scale. The HHIA differs from the HHIE in the content of some of the items. Replacement items on the HHIA focus on the occupational effects of hearing impairment. The APHAB was modified from the longer Profile of Hearing Aid Benefit (PHAB; Cox & Gilmore, 1990). The APHAB: (a) contains 24 items; (b) asks respondents to rate their experiences both with and without their hearing aids on a 7-point scale; and (c) produces a Global score and four sub-scales: Ease of Communication, Reverberation, Background Noise, and Aversiveness to sound.

Methods

The readability of each text (i.e., each self-report assessment tool) was assessed using the Windows-based software Reading Calculations version 7.5 (Micro Power and Light Co., 2008). This program is able to assess readability of documents using nine different, widely-used, and popular readability formulas. Unfortunately, there is no standard for choosing readability formulas (Breese & Burman, 2005). For this study, the FORCAST (Caylor, Sticht, Fox, & Ford, 1973), Flesch Reading Ease (FRE; Flesch, 1948), and Gunning FOG (Gunning, 1952) formulas were selected. Of the three, the FRE formula is considered to be the most widely- and heavily-used (Ley & Florio, 1996); however, the FORCAST formula is most appropriate for this study as it focuses on functional liter-
acy and is typically used in evaluating adult questionnaires, forms, lists, many websites, tests, and job materials not in narrative or prose form. What makes FORCAST different from other readability formulas is that it ignores the number of sentences, their lengths, and any hard punctuation (Redish, 2000), though many questionnaires do use complete sentences for items. The FORCAST formula calculates a reading grade level with lower grades (e.g., 1st grade) scored as easier to read and higher grades (e.g., 12th grade) as difficult to read. The FORCAST formula has a 0.66 correlation (DuBay, 2004) with performance on a standardized comprehension reading test (e.g., 42), and it has been demonstrated to correlate highly with other readability formulas (Caylor et al., 1973). For these reasons, the FRE and the FOG formulas were also considered for comparison. The Flesch-Kincaid (F-K) is a variant of the FRE formula and calculates a reading grade level, but it was designed more for technical information and legal documents such as insurance policies (Kincaid, Braby, & Mears, 1988; McClure, 1987).

The FRE formula is typically used with reading materials written for adults, it is considered to be the most widely used formula, and has been incorporated into the popular Microsoft® Word® software. Rather than reporting reading grade level, FRE scores range from 0 to 100 with lower scores indicating more difficult reading material. An advantage to using the FRE formula is that the score can be converted to an approximate reading grade level. One well-known limitation of the FRE score calculation by Microsoft® Word® is that it caps the highest grade level at the 12th grade; however, for this study, a different readability software package was used to allow for higher reading grade level calculations. The FRE formula has a 0.70 correlation with performance on a comprehension reading test (DuBay, 2004).

Unlike the FORCAST and FRE formulas, the FOG formula takes into account the number of hard words (two or more syllables). The FOG formula was designed for adult readers (Gunning, 1952) and for this reason has been popularized for use with healthcare and business reading materials. As with the FORCAST, the FOG formula also computes to a reading grade level. It also has a 0.91 correlation with performance on a comprehension reading test (DuBay, 2004). In a large scale readability study on anxiety and depression patient-reported outcome questionnaires by McHugh and Behar (2009), the FRE and FORCAST formulas were correlated strongly between 0.71 and 0.95. See the Appendix for the use, specific steps, and mathematical calculations applied by these formulas.

**Results**

Table 1 shows the means, medians, standard deviations, and ranges for each of the questionnaires produced by each of the formulas. An examination of the FORCAST calculations shows that the grade level of the four self-report assessment tools ranged from 7.8 to 11.2. The mean and the median of the four self-re-
port assessment tools were 9.5, with a standard deviation of 1.4. The results of the FOG calculations revealed a mean of 6.2, median of 5.8, and a standard deviation of 1.1 for the four self-report assessment tools. The results of the FRE calculations revealed a mean of 79.0 (equivalent to seventh-grade level), a median of 79.5 (equivalent to seventh-grade level), and a standard deviation of 4.7 for the four questionnaires. Table 2 presents the text-based quantitative variables used in the chosen formulas for each questionnaire. These variables include the number of words, the number of syllables, number of monosyllabic words, number of words with three or more syllables, number of hard words (FOG), and number of sentences.

<table>
<thead>
<tr>
<th>Questionnaire</th>
<th>FRE (grade level)</th>
<th>FOG</th>
<th>FORCAST</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHIE</td>
<td>78.0 (7th)</td>
<td>6.0</td>
<td>9.6</td>
</tr>
<tr>
<td>HAPI</td>
<td>73.0 (7th)</td>
<td>5.4</td>
<td>11.2</td>
</tr>
<tr>
<td>HHIA</td>
<td>81.0 (6th)</td>
<td>5.5</td>
<td>9.4</td>
</tr>
<tr>
<td>APHAB</td>
<td>84.0 (6th)</td>
<td>7.8</td>
<td>7.8</td>
</tr>
</tbody>
</table>

| Mean          | 79.0              | 6.2 | 9.5     |
| Median        | 79.5              | 5.8 | 9.5     |
| SD            | 4.7               | 1.1 | 1.4     |
| Range         | 73 to 82          | 5.4 to 8.7 | 7.8 to 11.2 |

Note. FRE = Flesch Reading Ease; FOG = Gunning Fog Index Readability Formula; FORCAST = FORCAST Readability Formula; HHIE = Hearing Handicap Inventory for the Elderly; HAPI = Hearing Aid Performance Inventory; HHIA = Hearing Handicap Inventory for Adults; APHAB = Abbreviated Profile of Hearing Aid Benefit.

*FRE conversion to grade level: 90-100 = 5th grade; 80-90 = 6th grade; 70-80 = 7th grade; 60-70 = 8th and 9th grade; 50-60 = 10th through 12th grades.

The results of this study show that the reading level of the four self-report assessment tools analysed in this study exceed the fifth-grade reading level recommended by experts in health literacy, regardless of the readability formulas used. When examining the results of the FORCAST formula, which is the most appropriate formula to use with questionnaires, differences were found among the four questionnaires. While the various reading formulas take many factors into account, generalisations can be made based on these results. Generally, the higher the proportion of monosyllabic words to overall words, the lower the reading grade level. Similarly, the lower the proportion of polysyllabic words (those containing three or more syllables) to overall words, the lower the reading grade level, resulting in improved readability.

DISCUSSION

The results of this study show that the reading level of the four self-report assessment tools analysed in this study exceed the fifth-grade reading level recommended by experts in health literacy, regardless of the readability formulas used. When examining the results of the FORCAST formula, which is the most appropriate formula to use with questionnaires, differences were found among the four questionnaires. While the various reading formulas take many factors into account, generalisations can be made based on these results. Generally, the higher the proportion of monosyllabic words to overall words, the lower the reading grade level. Similarly, the lower the proportion of polysyllabic words (those containing three or more syllables) to overall words, the lower the reading grade level, resulting in improved readability.
Table 2
Text-Based Quantitative Features of the Questionnaires

<table>
<thead>
<tr>
<th>Questionnaire</th>
<th>Words</th>
<th>Syllables(^a)</th>
<th>Monosyllabic words(^b)</th>
<th>Words of 3 or more syllables(^b)</th>
<th>Hard words (FOG)(^b)</th>
<th>Sentences(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHIE</td>
<td>436</td>
<td>625 (1.4)</td>
<td>301 (69%)</td>
<td>40 (9%)</td>
<td>31 (7%)</td>
<td>56 (7.8)</td>
</tr>
<tr>
<td>HAPI</td>
<td>2272</td>
<td>3499 (1.5)</td>
<td>1330 (59%)</td>
<td>225 (10%)</td>
<td>223 (10%)</td>
<td>624 (3.6)</td>
</tr>
<tr>
<td>HHIA</td>
<td>460</td>
<td>644 (1.4)</td>
<td>324 (70%)</td>
<td>36 (8%)</td>
<td>29 (6%)</td>
<td>61 (7.5)</td>
</tr>
<tr>
<td>APHAB</td>
<td>843</td>
<td>1106 (1.3)</td>
<td>688 (82%)</td>
<td>68 (8%)</td>
<td>67 (8%)</td>
<td>73 (11.5)</td>
</tr>
<tr>
<td>Mean</td>
<td>1002.8</td>
<td>1468.5</td>
<td>660.8</td>
<td>92.3</td>
<td>87.5</td>
<td>203.5</td>
</tr>
<tr>
<td>Median</td>
<td>651.5</td>
<td>875.0</td>
<td>506.0</td>
<td>54.0</td>
<td>49.0</td>
<td>67.0</td>
</tr>
<tr>
<td>SD</td>
<td>866.5</td>
<td>1371.8</td>
<td>480.1</td>
<td>89.6</td>
<td>92.0</td>
<td>280.4</td>
</tr>
</tbody>
</table>

Note. FOG = Gunning Fog Index Readability Formula; HHIE = Hearing Handicap Inventory for the Elderly; HAPI = Hearing Aid Performance Inventory; HHIA = Hearing Handicap Inventory for Adults; APHAB = Abbreviated Profile of Hearing Aid Benefit.

\(^a\)Parentheses indicate average number of syllables per word. \(^b\)Parentheses indicate percentage of words out of total number of words. \(^c\)Parentheses indicates average number of words per sentence.
It is important to keep in mind two main limitations of readability formulas. They do not provide information about the reader’s comprehension of the text. As stated above, readability is the culmination of several factors, including the characteristics of the reader. In addition, these formulas do not take those reader variables such as familiarity with the content of the questionnaire, the motivation of the reader, the cultural competency of the reader, the layout of the questionnaire, or the appropriateness of the language to the target audience (see Doak et al., 1996; Friedman & Hoffman-Goetz, 2006; Klare, 1963; Meade & Smith, 1991; Pichert & Elam, 1985). Therefore the readability data presented here provide an estimation of the readability of these self-report assessments. The readability may be higher or lower for individuals completing these assessments.

It is clear from these results, however, that the readability of commonly used self-report assessment tools in audiology exceed the recommended grade level proposed by health literacy advocates. Clinicians and researchers should be aware of the potential problems encountered by administering questionnaires with low readability to clients. One potential problem is that clients will not be able to answer the questions appropriately and the results from the assessment will not be valid. Another potential problem is that a client who completes a questionnaire at the beginning of a session, or in preparation for a session, will be put off by the lack of readability of the questionnaire. These problems in turn, could have a negative impact on the client-clinician relationship. That is, clients may enter the relationship with a negative view of themselves (because they found the questionnaire difficult to read) or of their clinicians (because they asked them to complete a questionnaire that was difficult to read). It is likely that the clinician will interpret negative attitudes from the client as a sign of resistance and will begin counseling with the client in terms of acceptance of hearing loss.

### Table 3
Practical Resources for Developing Easy-to-Read Health Materials

<table>
<thead>
<tr>
<th>Resource</th>
<th>Web Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear communication: An NIH literacy initiative</td>
<td><a href="http://www.nih.gov/clearcommunication/">http://www.nih.gov/clearcommunication/</a></td>
</tr>
<tr>
<td>Plain language: Improving communication from the federal government to the public</td>
<td><a href="http://www.plainlanguage.gov/">http://www.plainlanguage.gov/</a></td>
</tr>
</tbody>
</table>

It is important to keep in mind two main limitations of readability formulas. They do not provide information about the reader’s comprehension of the text. As stated above, readability is the culmination of several factors, including the characteristics of the reader. In addition, these formulas do not take those reader variables such as familiarity with the content of the questionnaire, the motivation of the reader, the cultural competency of the reader, the layout of the questionnaire, or the appropriateness of the language to the target audience (see Doak et al., 1996; Friedman & Hoffman-Goetz, 2006; Klare, 1963; Meade & Smith, 1991; Pichert & Elam, 1985). Therefore the readability data presented here provide an estimation of the readability of these self-report assessments. The readability may be higher or lower for individuals completing these assessments.

It is clear from these results, however, that the readability of commonly used self-report assessment tools in audiology exceed the recommended grade level proposed by health literacy advocates. Clinicians and researchers should be aware of the potential problems encountered by administering questionnaires with low readability to clients. One potential problem is that clients will not be able to answer the questions appropriately and the results from the assessment will not be valid. Another potential problem is that a client who completes a questionnaire at the beginning of a session, or in preparation for a session, will be put off by the lack of readability of the questionnaire. These problems in turn, could have a negative impact on the client-clinician relationship. That is, clients may enter the relationship with a negative view of themselves (because they found the questionnaire difficult to read) or of their clinicians (because they asked them to complete a questionnaire that was difficult to read). It is likely that the clinician will interpret negative attitudes from the client as a sign of resistance and will begin counseling with the client in terms of acceptance of hearing loss.
Another potential problem involves the validity of empirical data collected using questionnaires that have poor readability scores. That is, to what extent is a researcher measuring what is purported to be measured given that the outcome measure may not have a satisfactory readability score.

To resolve the potential problems associated with questionnaires with poor readability scores, researchers need to take readability into account when designing questionnaires for use in the clinical and research settings. To accomplish this goal, the reading level of the questionnaire should be as low as possible. It should use plain language and clear communication, paying particular attention to factors that improve readability such as clear instructions and questionnaire items with familiar words and preferably with words that do not have a high number of syllables. Table 3 offers practical resources for developing easy-to-read and easy-to-use health materials, including questionnaires. For those questionnaires that are already in common use, these could be revised by incorporating some of the resources’ tips and psychometrically re-analyzing them, making them more valid and ultimately producing more meaningful results in both clinical and research settings.

REFERENCES


**APPENDIX**

**READABILITY FORMULAS**

**The Flesch Reading Ease Readability Formula**

\[
RE = 206.835 - (1.015 \times ASL) - (84.6 \times ASW),
\]

where \(RE\) = Readability Ease; \(ASL\) = Average Sentence Length (i.e., the number of words divided by the number of sentences); and \(ASW\) = Average number of Syllables per Word (i.e., the number of syllables divided by the number of words). The output is a number ranging from 0 to 100. The higher the number, the easier the text is to read.

**The Gunning’s Fog Index (or FOG) Readability Formula**

\[
\text{Grade Level} = 0.4(\text{ASL} + \text{PHW}),
\]

where \(\text{ASL}\) = Average Sentence Length (i.e., number of words divided by the number of sentences) and \(\text{PHW}\) = Percentage of Hard Words. Short sentences written in Plain English achieve a better score than long sentences written in complicated language. Requires a minimum of 100 sample words.

**The FORCAST Readability Formula**

\[
\text{Grade Level} = 20 - (N/10),
\]

where \(N\) = Number of monosyllabic words in the sample text.