

Readability Assessment of Self-Report Hyperacusis Questionnaires

*Magdalena Margol-Gromada^a, Magda Sereda^{a, b}, David M. Baguley^{a, b, c}

^a Hearing Sciences, Division of Clinical Neuroscience, School of Medicine, University of Nottingham, Nottingham, UK; ^b NIHR Nottingham Biomedical Research Centre, Nottingham, UK; ^c Nottingham Audiology Services, Nottingham University NHS Trust, Nottingham, UK

Corresponding author: Magdalena Margol-Gromada, Hearing Sciences, Division of Clinical Neuroscience, School of Medicine, University of Nottingham, Building 40, University Park, Nottingham, NG7 2RD. E-mail: Magdalena.Margol-Gromada@nottingham.ac.uk.

Abstract

Objective: To assess the overall readability of five currently available hyperacusis questionnaires and to assess the variability of single items within each questionnaire. **Design:** Comparative study of self-report hyperacusis questionnaires: (1) Geräuschüberempfindlichkeits-Fragebogen (GUF), (2) Noise Avoidance Questionnaire (NAQ), (3) Hyperacusis Questionnaire (HQ), (4) Sound Sensitive-Tinnitus Index (SSTI), and (5) Inventory of Hyperacusis Symptoms (IHS). Well-established readability formulas Flesh-Kincaid Grade Level (FKGL), Flesch Reading Ease (FRE), Simple Measure of Gobbledygook (SMOG) and FORCAST and a computerised readability calculation software were used. **Study sample:** Five questionnaires. **Results:** Reading levels calculated by each formula varied for every questionnaire. Readability scores ranged from 7.7th to 12.7th grade for overall readability depending on the questionnaire. This exceeded the grade reading levels of 5th to 6th grade (10-12 years old) as recommended by the American Medical Association or 7th to 8th grade (12-14 years old) as recommended by the US National Institutes of Health. Single item readability analysis based on FKGL revealed that 32% to 70% of single items are written above the recommended grade levels. **Conclusion:** All five questionnaires are written at close to or exceeding the recommended grade levels. This requires attention from developers but also when interpreting the questionnaire scores obtained in clinic.

Keywords: readability, health literacy, hyperacusis, self-report questionnaires, reduced sound tolerance

Introduction

Hyperacusis is an emergent diagnosis and a growing field of interest in both the clinical and research communities. It has been characterised as the ‘perception of everyday environmental sounds as being overwhelmingly loud or intense’ (Fackrell et al. 2017) and as Aazh et al. (2016) notes the impacts of this span social, professional and recreational contexts. Tyler et al. (2014) suggest that there are sub-types of hyperacusis, using a classification system of whether pain, fear, annoyance or loudness is the defining feature of one’s experience.

Commonly co-incident with tinnitus, hyperacusis has often been assessed and measured using tinnitus-specific questionnaires, and were first linked by (Tyler and Conrad-Armes 1983). Understanding and knowledge of the condition has improved and there are now hyperacusis specific tools for diagnostic and measurement purposes. However, they are limited in number and vary in degrees of robustness and psychometric validation (Fackrell and Hoare 2018). In brief, the questionnaires are concerned with gathering information and examples of when and how hyperacusis affects the individual and to what extent, by grading the severity of hyperacusis based on the patient’s answers.

Questionnaires are a fundamental part of clinical practice, particularly when concerned with subjective, self-reported symptoms and conditions such as hyperacusis. They allow for a time-efficient and structured assessment of symptoms and experience as well as facilitating the assessment of changes over time (Douglas and Kelly-Campbell 2018). Specifically for people with hyperacusis seeking clinical intervention, questionnaires can be a tool for articulating the diverse and challenging symptoms they experience with a consistent style of questioning. This is especially important as collecting data on subjective symptoms such as hyperacusis, is strongly influenced by how the question about the experience is formulated (Baguley and Hoare 2018). Furthermore, people with hyperacusis can experience symptoms such as fatigue and concentration difficulties in situations that are well tolerated by other

members of the population (Paulin, Andersson, and Nordin 2016), for example the Audiology clinic, which may interfere with one's ability to read, comprehend and answer questions. Therefore, it is clinically important to ensure that patients can easily read and understand the questionnaires used.

The readability of a particular text is the objective measure of the reading skills the person should have in order to be able to understand it (Badarudeen and Sabharwal 2010). This is quantified as the number of years of education equivalent to a reading grade level in the US grade system, but can also be converted for other country specific systems, e.g. the key stage system in the UK. Commonly used formulas are the Flesch Reading Ease (FRE) and Flesch-Kincaid Grade Level (FKGL) (Flesch 1948). The Simple Measure of Gobbledygook (SMOG) (McLaughlin 1969) is recommended for use with healthcare materials as it is based on more recent criteria for determining reading grade level, and has been reported to be the most suited and practical for application to health care materials (Wang et al. 2013). These formulas take into consideration number of syllables per word and/ or average words per sentence and are typically meant for use with prose-like text. This is problematic as questionnaires are rarely written in prose-like form, and are more typical to have a disjointed or stem and leaf format. The FORCAST formula (Caylor et al. 1973) is deemed to be most suited to assessing readability of text not in prose-like form such as questionnaires, forms, lists test and job materials (Atcherson et al. 2013). It does not count number of sentences, or their average length but rather counts the number of monosyllabic words.

Single item analysis is another area of difficulty, as currently there are no widely used and validated readability formulas developed for single-item analysis specifically. Calderón et al. (2006) applied the FRE and FKGL on single items comprising popular Quality of Life Questionnaires, after combining stem-leaf format questions to form full sentences, in order to comply with recommendations to only assess running text (Flesch 1979). This method was

also used by Betschart et al. (2018), and appears to be successful in overcoming the methodological challenge of assessing single items and text not in prose-like form.

Readability is an integral part of health literacy, definitions of which vary in the literature. A systematic review by Sørensen et al. (2012) arrived at the definition below, following the thematic analysis of 17 eligible publications:

'Health literacy is linked to literacy and entails people's knowledge, motivation and competences to access, understand, appraise, and apply health information in order to make judgments and take decisions in everyday life concerning healthcare, disease prevention and health promotion to maintain or improve quality of life during the life course.'

Despite the importance of health literacy it is commonly neglected, a concerning issue given that health literacy is the single best predictor of an individual's health status (Badarudeen and Sabharwal 2010). Low literacy is associated with severe adverse health outcomes, including increased incidence of chronic illness and poor use of preventative health services (Berkman et al. 2011). In the US over \$230 billion a year is linked to low adult literacy, with nearly 50% of Americans finding understanding and using health information difficult (ProLiteracy 2019). In the UK, current health information is written at a level too complex for 43% of adults aged 16-65, if numeracy skills are required for comprehension of the information the figure rises to 61% (Rowlands et al. 2015). Ensuring that patients are able to fully understand matters pertaining to their health is a fundamental part of good practice. Therefore, adult health materials should be prepared at the lowest possible level of reading difficulty which is generally 5th grade (Weiss and Coyne 1997). More specifically Gilligan and Weinstein (2014) have reported important links between health literacy and rehabilitation outcomes for patients in the audiology clinic. The authors argue that in order for patient-centred care to succeed we need to ensure that the tools we use to enable the patient perspective

within management options are suitable i.e. readable by their target audience. If the readability of a questionnaire is too difficult, this can lead to patients rejecting the questionnaire, providing partial information or answering in a way that does not truly reflect their experience (Atcherson et al. 2013). Therefore, reading grade levels of 5th to 6th grade (10-12 years old) are recommended by the American Medical Association (AMA) (Weiss 2007) and 7th to 8th grade (12-14 years old) are recommended by the U.S. National Institutes of Health (NIH) (Medlineplus 2017). Studies reporting on readability for tinnitus questionnaires (Atcherson, Zraick, and Brasseux 2011), Auditory Processing Disorder questionnaires (Atcherson et al. 2013), and Adult Audiology Rehabilitation outcome measures (Douglas and Kelly-Campbell 2018) indicate that readability of questionnaires and patient-reported outcome measures generally exceed the recommended reading levels mentioned above. The above-mentioned studies evaluated the overall readability of questionnaires meaning that information about the variability in the readability of single items was not addressed. This is an important omission as very easy to read items can potentially skew the overall readability of the questionnaire towards a lower grade, hiding the more difficult items. The resulting readability level would not reflect the true difficulty of the text (Homan, Hewitt, and Linder 1994, Betschart et al. 2018). Unlike running prose in which context can help the reader comprehend the meaning of a given text, respondents are required and expected to comprehend each item in a questionnaire separately (Calderón et al. 2006). Therefore, analysis of single items allows for a more comprehensive assessment of readability and can highlight particular items within a questionnaire that require caution when interpreting patient's answers (Betschart et al. 2018).

Pertinent to the present article is the issue surrounding the readability of translated text, because when not performed robustly, translation can lead to semantic, contextual and

cultural differences that ultimately change the meaning of the questions and therefore the potential responses to the questions (Hall et al. 2018). Thorough guidelines on translation of questionnaires for different languages and cultures do exist, and state that the process should be systematic in its approach and include forward and backward translation, mono- and bilingual testing, and consultation with experts in the field (Hall et al. 2018, Maneesriwongul and Dixon 2004). The GUF, HQ and NAQ were not developed in the English language, and the translation process is not well detailed in the source publications. At present, there does not seem to be any readability analysis of the GUF, HQ and NAQ in the source language, however there are indications that if the source language and the translated language are from the same language family the readability level of the text is similar (Coco et al. 2017, Ciobanu, Dinu, and Pepelea 2015). Furthermore, these questionnaires are used or intended for use in the clinic and for research purposes (see (Aazh, Lammaing, and Moore 2017, Aazh and Moore 2017); hence, there is a need to be aware of the considerations that need to be made when interpreting the answers of these questionnaires. The purpose of the present study is to assess the overall readability of five currently available self-report hyperacusis questionnaires and to assess the variability in readability of single items within each questionnaire.

Methods

Questionnaire selection

Questionnaires were selected based on the following criteria: (i) focus on quantifying and characterising an individual's sound tolerance difficulties, (ii) questionnaire designed to be completed by the patient without help or guidance from a clinician, (iii) questionnaires used or intended for use in the clinic and (iv) questionnaires that have undergone psychometric validation. Questionnaires that were designed to be administered as a part of a semi-structured interview, or with clinician involvement were excluded.

Five questionnaires meeting the above criteria were selected for analysis:

- Geräuschüberempfindlichkeits-Fragebogen (GUF); English translation (Nelting and Finlayson 2004)
- Noise Avoidance Questionnaire (NAQ); English translation (Blaesing and Kroener-Herwig 2012)
- Hyperacusis Questionnaire (HQ); English translation (Khalifa et al. 2002)
- Sound Sensitive-Tinnitus Index (SSTI) (Greenberg 2018)
- Inventory of Hyperacusis Symptoms (IHS) (Greenberg and Carlos 2018)

Formula selection

There are a number of readability formulas in existence; however, no widely used formulas that are specifically designed for single-item analysis. We selected the FRE and the FKGL as they are most frequently utilised in readability literature facilitating comparison between the present study and published literature; the FORCAST as it is deemed the most suitable for readability analysis of questionnaires; and the SMOG formula as it has been reported to be most consistent, and most appropriate for application to healthcare material given it has an expected comprehension of 100% (Wang et al. 2013).

The FKGL, FORCAST and SMOG formulas give reading grade level scores, the higher the grade the more difficult the text is. The FRE is scored on a scale from 0-100, where a lower number is equivalent to a higher grade level score.

English versions of the questionnaires were first copied into a Word document, copying the exact format that was found in the source by author MMG. Second author MS then checked the copied versions for accuracy against the original source. To assess the overall readability of each questionnaire full sentences were formed from preamble statements and question options; stem and leaf format questions were combined to form full sentences for each option. Question options that included short repeated words e.g. the four point scale 'never, sometimes, often, very often', are likely to score very easy and skew results therefore were removed from the analysis. This approach has been utilised by Calderón et al. (2006) and later by Betschart et al. (2018), based on the recommendation that the FRE should only be used to test running text (Flesch 1979). All additional text including references and notes to the clinician were excluded from the analysis.

Readability analysis

Readability analysis was conducted using the software package Readability Studio Professional Edition version 2015 for Windows, (Oleander Software, Ltd, Vandalia, OH, USA). Descriptive statistics for mean, median, range and standard deviation were calculated using Microsoft Excel 2016 for Windows 10.

The readability analysis was conducted in three parts:

- (1) Readability assessment of the questionnaires in their original format using the FORCAST formula.

- (2) Readability assessment of the questionnaires where each item was manipulated to form full sentences, using the FRE, FKGL and SMOG formulas.
- (3) Readability assessment of single items comprising each questionnaire, using the FRE and FKGL formulas, following the approach utilised by Calderón et al. (2006).

Results

Factors that may affect readability such as number of words with more than three syllables and sentence length were extracted (Table 1), and can be useful when considering changes to the text to reduce the reading grade level. However, it is important to note that readability formulas in general, and those utilised in the present study, do not always use the same factors in their mathematical formula to determine the grade reading level score for a particular text.

Table 1 goes here.

Readability analysis using the FORCAST formula

Readability analysis of the questionnaires in the original formats using the FORCAST formula showed that the overall readability grade level for each of the five questionnaires exceeded the recommended reading level of 5th to 6th (10-12 years old) and 7th to 8th grade (12-14 years old), see Table 2. The software reports results according to the US grade system, which can be converted to country specific school years, in this case the UK year-group system can be inferred by adding one year to the US grade results.

Table 2 goes here.

Readability analysis using the FRE, FKGL and SMOG formulas

The grade levels as calculated by the FRE, FKGL and SMOG formulas for the questionnaires in the manipulated, full sentence format are shown in Table 2. There are some evident differences in the reading grade levels for each questionnaire depending on the readability formula used, which is to be expected given that each formula takes into consideration slightly different factors. With reference to the more conservative recommendations, all questionnaires exceed the grade 5th to 6th reading level. However, four out of the five questionnaires fall within the 7th to 8th grade level according to the FRE, and three out of five fall within the above criteria according to the FKGL. The SSTI has the highest reading grade levels as determined by each of the three formulas thus requiring a more advanced reading age than compared to the other questionnaires. SMOG reading grade levels for the IHS and the HQ exceed both recommendations; however fall within the recommendations according to the FKGL formula. The differences in reading grade levels could be attributed to differences between the formulas, although notably the SMOG formula yielded higher-grade reading levels for all questionnaires than compared to FRE and FKGL.

The SSTI reading grade level as calculated by the FRE is equivalent to that of a 10th-12th grader (15-18 years old), as it has a high average number of words per sentence and average syllables per word. It is important to note that the FRE assumes that plain English has a score between 60-70, approximately equivalent to 7th grade level, however the recommendations for healthcare materials is even lower, as above.

Readability analysis of single items using FRE and FGKL formulas

Single item analysis revealed variability in single item readability within each questionnaire assessed using the FKGL (Figure 1). Results show that a readability level above the maximum recommended 8th grade level was found for 47% of items in the GUF, 44% of items in the NAQ, 39% of items in the HQ, 70% of items in the SSTI, and 32% of items in the IHS. The highest score was a grade level of 16 according to the FKGL, suggesting college level educational attainment would be required to read this with comprehension, and was found for single items within the NAQ, SSTI, and IHS. According to the FRE, scores of 6, 15 and 27 were found for single items in the IHS, SSTI and NAQ respectively (a lower score indicates more difficult readability), meaning that a person would need the reading ability of a postgraduate to understand them.

Figure 1 here.

Discussion

The present study assessed the readability of the currently available self-report hyperacusis questionnaires. Results revealed a range of readability grade levels for each questionnaire and across formulas, most of which clearly exceeded the recommended grade reading levels of 5th to 6th grade and 7th to 8th grade endorsed by the AMA and the NIH, respectively.

Readability analysis using the FORCAST formula

All questionnaires exceeded the grade reading level recommendations as per the FORCAST formula. Similar to these findings are reports from Atcherson, Zraick, and Brasseur (2011) who used FORCAST to assess tinnitus questionnaire readability. However, the expected comprehension for the FORCAST formula is only 35% and so it may not be the most appropriate formula to use for the healthcare setting. Ensuring that patients can comprehend 100% of the information they receive relating to their healthcare is an integral part of healthcare literacy and facilitates better healthcare outcomes for patients (Gilligan and Weinstein 2014, Douglas and Kelly-Campbell 2018). As patient-reported outcome measures are increasingly employed in clinical practice, research and used to inform healthcare services, it is crucial that patients are able to read and understand the questions (El-Daly et al. 2016). Furthermore, the NAQ yielded the highest grade reading level with the FORCAST of 12.7th grade. Assuming the average reading age of UK adults to be approximately that of a 5th to 6th grader (10-11 year old) this would mean that around 5 million adults would not be able to read and comprehend the questionnaires (National Literacy Trust 2017). In the US this would translate to approximately 30 million adults that are classified as having a below basic health literacy level and not being able to comprehend the questionnaires (U.S. Department of Health and Human Services 2008). Of the individuals that could read it, they still may not

comprehend 100% of the questionnaire, meaning that in this case only someone with a more advanced reading age would be able to achieve 100% comprehension.

Readability analysis using the FRE, FGKL and SMOG formulas

A common report in the readability literature is around the variability in the grade levels given by different formulas, which can be attributed to the parameters analysed by each formula (Atcherson et al. 2013). Similarly, in the present study, grade levels for a particular questionnaire varied by approximately three grades depending on the formula (Table 2). SMOG reading grade levels for every questionnaire clearly exceeded even the less conservative recommendation of 7th to 8th grade reading level, and whilst correlation analyses were not conducted the SMOG and FKGL generally showed the same trend. Similar results have been reported for audiological rehabilitation outcome measures by Douglas and Kelly-Campbell (2018), with the SMOG and FORCAST formulas yielding the highest reading grade level. The SMOG formula is not only recommended for use with healthcare materials but is based on more recent criteria for determining readability (Wang et al. 2013).

A difficulty that is not often discussed in the literature concerned with readability of questionnaires is that fact that other than the FORCAST, the other formulas should be applied to text in full sentences. This is why it was necessary to carry out the manipulation of the questionnaires to achieve the most appropriate format for the SMOG, FRE, and FKGL formulas. This highlights the strong need for a formula that can meet the specific requirements of 100% comprehension as well as being able to tackle non-prose like, stem and leaf formats questionnaires.

Readability analysis of single items using the FRE and FGKL formulas

Readability analysis of single-items allowed further insight into the readability issues within each questionnaire. The SSTI single item scores clearly exceeded the other four questionnaires, with 70% of single items yielding a grade level above the maximum 8th grade recommendation. This could affect the clinical usefulness of information gathered with these questionnaires, as patients with lower reading ages would be at risk of rejecting the questionnaires, or providing inaccurate information (Atcherson, Zraick, and Brasseur 2011). Whilst there are no guidelines on single item readability, reports in the literature suggest that the best way to facilitate meaningful, reliable and useful information gathering using questionnaires as a tool would be to ensure that every item is readable by patients (Gilligan and Weinstein 2014).

Readability is part of the wider concept of health literacy; poorer health literacy is associated with poorer healthcare outcomes. Douglas and Kelly-Campbell (2018) argue that ignoring basics such as the readability of self-report patient materials can lead to patients' issues not being fully addressed. For example, if a patient has misunderstood questions or answered inaccurately this may result in a contraindicated early discharge. In research, the issue may reduce the accuracy of evaluating interventions, if the participant has not been able to understand a patient-report tool used as the outcome measure (Douglas and Kelly-Campbell 2018) potentially invalidating the empirical data collected (Atcherson, Zraick, and Brasseur 2011). Apart from potentially affecting the validity of the information gathered, the use of patient materials that are not suitable for even those with the lowest health literacy can create barriers to patients in accessing services they need (Rajah et al. 2018). There may be additional factors to be aware of, especially for patients with hyperacusis. It is common to have patients fill in questionnaires as they wait for appointments. Clinic waiting rooms are known to be busy and commonly noisy environments that can induce stress in patients even without

hyperacusis, especially if noise levels exceed ambient noise level recommendations (Hill and LaVela 2015). A difficult soundscape coupled with a questionnaire that is written at a reading level exceeding that of the patient, could introduce further difficulties for patients with hyperacusis.

Recommendations

Improving the readability of single items and questionnaires as a whole should be a consideration for developers of questionnaires. Readability can be improved (made easier) by using mono- or bi-syllabic word substitutions for unfamiliar medical terms, using shorter sentences, avoiding use of jargon and using the active voice (Weiss 2007, El-Daly et al. 2016). Furthermore, it is important to avoid including too much content within a single-item and sticking to one or two key topics to address (Weiss 2007). Edited versions of questionnaires should undergo re-validation.

Limitations

A limitation of the present study, as with other reports, is that only the text was analysed for readability, whereas, readability is also affected by other factors such as format, text typeface and size and images (Atcherson, Zraick, and Brasseur 2011). Only English versions of the questionnaires were analysed, hence the reported results cannot be generalised to other translations that exist. Readability analysis on the original German language versions of the GUF, HQ and NAQ may be of clinical interest.

Despite the limitations, the present study presents important information on the readability of hyperacusis questionnaires, utilising format appropriate formulas and providing an insight into the single-item variability within some of the currently available questionnaires.

Conclusion

Researchers and developers should consider the overall readability of the developed questionnaires, ensuring that it is in keeping with recommendations. Furthermore, a greater awareness of and adherence to the recommendations should be made on the single item level, so that there is less variability within a questionnaire. Researchers and developers should exercise caution when interpreting patient responses to questions or survey items that require reading levels exceeding the published standards. There is a clear need for more work on methodological approaches to single-item readability, including the need for a formula that is specific for use with short samples of text, to allow for more robust single item analysis.

Acknowledgements

The views expressed are those of author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Declaration of interest

None declared.

ORCiDs

Magdalena Margol-Gromada: <https://orcid.org/0000-0002-2882-0786>

Magdalena Sereda: <https://orcid.org/0000-0002-2792-8496>

David Baguley: <https://orcid.org/0000-0002-0767-0723>

Social media

Twitter

Magdalena Margol-Gromada: @Magdalena_UoN

Magdalena Sereda: @MagdalenaSereda

David Baguley: @DrDavidBaguley

References

- Aazh, Hashir, Karen Lammaing, and Brian C. Moore. 2017. "Factors related to tinnitus and hyperacusis handicap in older people." *International Journal of Audiology* 56 (9):677-684.
- Aazh, Hashir, and Brian C. J. Moore. 2017. "Factors Associated With Depression in Patients With Tinnitus and Hyperacusis." *American Journal of Audiology* 26 (4):562-569. doi: 10.1044/2017_AJA-17-0008.
- Aazh, Hashir, Brian C. Moore, Karen Lammaing, and Mark Cropley. 2016. "Tinnitus and hyperacusis therapy in a UK National Health Service audiology department: Patients' evaluations of the effectiveness of treatments." *International Journal of Audiology* 55 (9):514-522.
- Atcherson, Samuel R., Cynthia M. Richburg, Richard I. Zraick, and Cassandra M. George. 2013. "Readability of Questionnaires Assessing Listening Difficulties Associated with (Central) Auditory Processing Disorders." *Language, Speech, and Hearing Services in Schools* 44 (1):48-60. doi: 10.1044/0161-1461(2012/11-0055).
- Atcherson, Samuel R., Richard I. Zraick, and Raven E. Brasseur. 2011. "Readability of Patient-Reported Outcome Questionnaires For Use With Persons With Tinnitus." *Ear and Hearing* 32 (5):671-673. doi: 10.1097/AUD.0b013e3182134654.
- Badarudeen, Sameer, and Sanjeev Sabharwal. 2010. "Assessing readability of patient education materials: current role in orthopaedics." *Clinical orthopaedics and related research* 468 (10):2572-2580. doi: 10.1007/s11999-010-1380-y.
- Baguley, D. M., and D. J. Hoare. 2018. "Hyperacusis: major research questions." *HNO* 66 (5):358-363. doi: 10.1007/s00106-017-0464-3.
- Berkman, N. D., S. L. Sheridan, K. E. Donahue, D. J. Halpern, and K. Crotty. 2011. "Low health literacy and health outcomes: an updated systematic review." *Ann Intern Med* 155 (2):97-107. doi: 10.7326/0003-4819-155-2-201107190-00005.
- Betschart, Patrick, Dominik Abt, Hans-Peter Schmid, Pascal Viktorin, Janine Langenauer, and Valentin Zumstein. 2018. "Readability assessment of commonly used urological questionnaires." *Investigative and clinical urology* 59 (5):297-304. doi: 10.4111/icu.2018.59.5.297.
- Blaesing, Lena, and Birgit Kroener-Herwig. 2012. "Self-reported and behavioral sound avoidance in tinnitus and hyperacusis subjects, and association with anxiety ratings." *International Journal of Audiology* 51 (8):611-617. doi: 10.3109/14992027.2012.664290.
- Calderón, José L., Leo S. Morales, Honghu Liu, and Ron D. Hays. 2006. "Variation in the readability of items within surveys." *American journal of medical quality : the official journal of the American College of Medical Quality* 21 (1):49-56. doi: 10.1177/1062860605283572.
- Caylor, J.S., T.G. Stitch, L.C. Fox, and J.P. Ford. 1973. "Development of a simple readability index for job reading material." Paper presented at the meeting the annual meeting of the American Educational Research Association, New Orleans, LA.
- Ciobanu, Alina Maria, Liviu P Dinu, and Flaviu Ioan Pepelea. 2015. "Readability Assessment of Translated Texts." *Natural Language Processing Hissar, Bulgaria*.
- Coco, Laura, Sonia Colina, Samuel R. Atcherson, and Nicole Marrone. 2017. "Readability Level of Spanish-Language Patient-Reported Outcome Measures in Audiology and Otolaryngology." *American journal of audiology* 26 (3):309-317. doi: 10.1044/2017_AJA-17-0018.

- Douglas, A., and R. J. Kelly-Campbell. 2018. "Readability of patient-reported outcome measures in adult audiologic rehabilitation." *American Journal of Audiology* 27 (2):208-218. doi: 10.1044/2018_AJA-17-0095.
- El-Daly, Ibraheim, Hajir Ibraheim, Karthig Rajakulendran, Paul Culpan, and Peter Bates. 2016. "Are patient-reported outcome measures in orthopaedics easily read by patients?" *Clinical orthopaedics and related research* 474 (1):246-255. doi: 10.1007/s11999-015-4595-0.
- Fackrell, Kathryn, and Derek J. Hoare. 2018. "Scales and questionnaires for decreased sound tolerance." In *Hyperacusis and disorders of sound intolerance: Clinical and research perspectives*, 43-58. San Diego, CA: Plural Publishing; US.
- Fackrell, Kathryn, Iskra Potgieter, Giriraj S. Shekhawat, David M. Baguley, Magdalena Sereda, and Derek J. Hoare. 2017. "Clinical Interventions for Hyperacusis in Adults: A Scoping Review to Assess the Current Position and Determine Priorities for Research." *BioMed research international* 2017:2723715-2723715. doi: 10.1155/2017/2723715.
- Flesch, Rudolf. 1948. "A new readability yardstick." *Journal of Applied Psychology* 32 (3):221-233. doi: 10.1037/h0057532.
- Flesch, Rudolf. 1979. "How to write plain English." accessed 10th October 2018. <http://pages.stern.nyu.edu/~wstarbuc/Writing/Flesch.htm>.
- Gilligan, Jennifer, and Barbara E Weinstein. 2014. "Health Literacy and Patient-Centered Care in Audiology – Implications for Adult Aural Rehabilitation." *Journal of Communication Disorders, Deaf Studies & Hearing Aids* 2 (2):4. doi: 10.4172/jcdsha.1000110.
- Greenberg, Benjamin. 2018. "Assessing the impact of sound sensitivity in tinnitus." *Dissertation Abstracts International: Section B: The Sciences and Engineering* 79 (1-B(E)):No Pagination Specified.
- Greenberg, Benjamin, and Megan Carlos. 2018. "Psychometric Properties and Factor Structure of a New Scale to Measure Hyperacusis: Introducing the Inventory of Hyperacusis Symptoms." *Ear And Hearing* 39 (5):1025-1034. doi: 10.1097/AUD.0000000000000583.
- Hall, Deborah A., Silvia Zaragoza Domingo, Leila Z. Hamdache, Vinaya Manchaiah, Spoorthi Thammaiah, Chris Evans, and Lena L. N. Wong. 2018. "A good practice guide for translating and adapting hearing-related questionnaires for different languages and cultures." *International Journal of Audiology* 57 (3):161-175. doi: 10.1080/14992027.2017.1393565.
- Hill, Jennifer N., and Sherri L. LaVela. 2015. "Noise Levels in Patient Rooms and at Nursing Stations at Three VA Medical Centers." *HERD: Health Environments Research & Design Journal* 9 (1):54-63. doi: 10.1177/1937586715592635.
- Homan, Susan, Margaret Hewitt, and Jean Linder. 1994. "The Development and Validation of a Formula for Measuring Single-Sentence Test Item Readability." *Journal of Educational Measurement* 31 (4):349-358. doi: 10.1111/j.1745-3984.1994.tb00452.x.
- Khalifa, S., S. Dubal, E. Veuillet, F. Perez-Diaz, R. Jouvent, and L. Collet. 2002. "Psychometric Normalization of a Hyperacusis Questionnaire." *ORL* 64 (6):436-442. doi: 10.1159/000067570.
- Maneesriwongul, Wantana, and Jane K. Dixon. 2004. "Instrument translation process: a methods review." *Journal of Advanced Nursing* 48 (2):175-186. doi: 10.1111/j.1365-2648.2004.03185.x.
- McLaughlin, G. H. 1969. "SMOG grading: A new readability formula." *Journal of Reading* 12 (8):639-646.

- Medlineplus. 2017. "How to Write Easy-to-Read Health Materials." U.S. National Library of Medicine, U.S. Department of Health and Human Services and National Institutes of Health, Last Modified 28 June 2017, accessed 11th November.
<https://medlineplus.gov/etr.html>.
- National Literacy Trust. 2017. "Adult literacy." accessed 22nd October.
<https://literacytrust.org.uk/parents-and-families/adult-literacy/>.
- Nelting, M, and N.K. Finlayson. 2004. *GÜF-Geräuschüberempfindlichkeits-Fragebogen*. Göttingen: Hogrefe.
- Paulin, Johan, Linus Andersson, and Steven Nordin. 2016. "Characteristics of hyperacusis in the general population." *Noise & Health* 18 (83):178-184. doi: 10.4103/1463-1741.189244.
- ProLiteracy. 2019. "Adult Literacy Facts." accessed 2nd January.
<https://proliteracy.org/Resources/Adult-Literacy-Facts>.
- Rajah, Retha, Mohamed Azmi Ahmad Hassali, Lim Ching Jou, and Muthu Kumar Murugiah. 2018. "The perspective of healthcare providers and patients on health literacy: a systematic review of the quantitative and qualitative studies." *Perspectives in Public Health* 138 (2):122-132. doi: 10.1177/1757913917733775.
- Rowlands, Gillian, Joanne Protheroe, John Winkley, Marty Richardson, Paul T Seed, and Rima Rudd. 2015. "A mismatch between population health literacy and the complexity of health information: an observational study." *British Journal of General Practice* 65 (635):e379-e386. doi: 10.3399/bjgp15X685285.
- Sørensen, Kristine, Stephan Van den Broucke, James Fullam, Gerardine Doyle, Jürgen Pelikan, Zofia Slonska, Helmut Brand, and European Consortium Health Literacy Project. 2012. "Health literacy and public health: A systematic review and integration of definitions and models." *BMC Public Health* 12 (1):80. doi: 10.1186/1471-2458-12-80.
- Tyler, R. S., and D. Conrad-Armes. 1983. "The determination of tinnitus loudness considering the effects of recruitment." *J Speech Hear Res* 26 (1):59-72. doi: 10.1044/jshr.2601.59.
- Tyler, Richard S., Martin Pienkowski, Eveling Rojas Roncancio, Hyung Jin Jun, Tom Brozoski, Nicolas Dauman, Claudia Barros Coelho, Gerhard Andersson, Andrew J. Keiner, Anthony T. Cacace, Nora Martin, and Brian C. J. Moore. 2014. "A Review of Hyperacusis and Future Directions: Part I. Definitions and Manifestations." *American Journal of Audiology* 23 (4):402-419. doi: doi:10.1044/2014_AJA-14-0010.
- U.S. Department of Health and Human Services. 2008. Quick Guide to Health Literacy. edited by Office of Disease Prevention and Health Promotion. Washington, DC: Washington, DC.
- Wang, Lih-Wern, Michael J. Miller, Michael R. Schmitt, and Frances K. Wen. 2013. "Assessing readability formula differences with written health information materials: Application, results, and recommendations." *Research in Social and Administrative Pharmacy* 9 (5):503-516. doi: 10.1016/j.sapharm.2012.05.009.
- Weiss, Barry, D. 2007. *Health literacy and patient safety: Help patients understand. Manual for clinicians*. 2nd ed. Chicago: American Medical Association and American Medical Foundation.
- Weiss, Barry, D., and Cathy Coyne. 1997. "Communicating with patients who cannot read." *The New England Journal of Medicine* 337 (4):272-274.

Table 1. Factors affecting readability of written text, for each questionnaire.

Questionnaire	No. of items	Total words	Complex words (3+syllables)	Long words (6+ characters)	Average sentence length (words)	Number of difficult sentences (more than 22 words)
GUF	15	374	21 (5.6%)	148 (39.6%)	12.5	0
NAQ	25	542	54 (10%)	166 (30.6%)	8.6	0
HQ	18	301	40 (13.3%)	86 (28.6%)	6.4	0
SSTI	20	712	56 (7.9%)	160 (22.5%)	17.7	3
IHS	25	429	31 (7.2%)	115 (26.8%)	10.2	0
Median	20	429	40	148	10.2	0
Mean	19.8	474	40.4	135	11.08	0.6
SD	5.26	157.5	14.94	33.75	4.32	1.34
Range	14-25	313-712	21-56	86-166	6.4-17.7	0-3

GUF- Geräuschüberempfindlichkeits-Fragebogen, NAQ- Noise Avoidance Questionnaire, HQ- Hyperacusis Questionnaire, SSTI- Sound Sensitive Tinnitus Index, IHS- Inventory of Hyperacusis Symptoms.

Table 2. FORCAST, FKGL, SMOG and FRE reading grade levels

Questionnaire	FORCAST	FKGL	SMOG	*FRE (grade level)
GUF	12.1	8.3	10.0	62 (8 th -9 th)
NAQ	12.7	8.9	9.6	68 (8 th -9 th)
HQ	9.9	7.8	10.8	63 (8 th -9 th)
SSTI	8.9	10.1	12.7	58 (10 th -12 th)
IHS	9.9	7.7	11.1	65 (10 th -12 th)
<i>median</i>	9.9	8.3	10.8	63 (8 th -9 th)
<i>mean</i>	10.7	8.56	10.84	63.2 (8 th -9 th)
<i>SD</i>	1.62	0.98	1.2	3.7

GUF- Geräuschüberempfindlichkeits-Fragebogen, NAQ- Noise Avoidance Questionnaire, HQ- Hyperacusis Questionnaire, SSTI- Sound Sensitive Tinnitus Index, IHS- Inventory of Hyperacusis Symptoms. Note that where a score has an integer after a decimal point e.g. 12.1 this means that the reading grade level is 12th grade, 1st month of academic year. FORCAST reading grade levels are for questionnaires in their original format. FRE, FKGL and SMOG scores are for the questionnaires in the manipulated full-sentence format. *Conversion from FRE to grade level: 90-100- 5th grade; 80-90- 6th grade; 70 to 80- 7th grade; 60-70- 8th and 9th grade; 50-60- 10th-12th grade; 30-50- college; 0 to 30- college graduate.

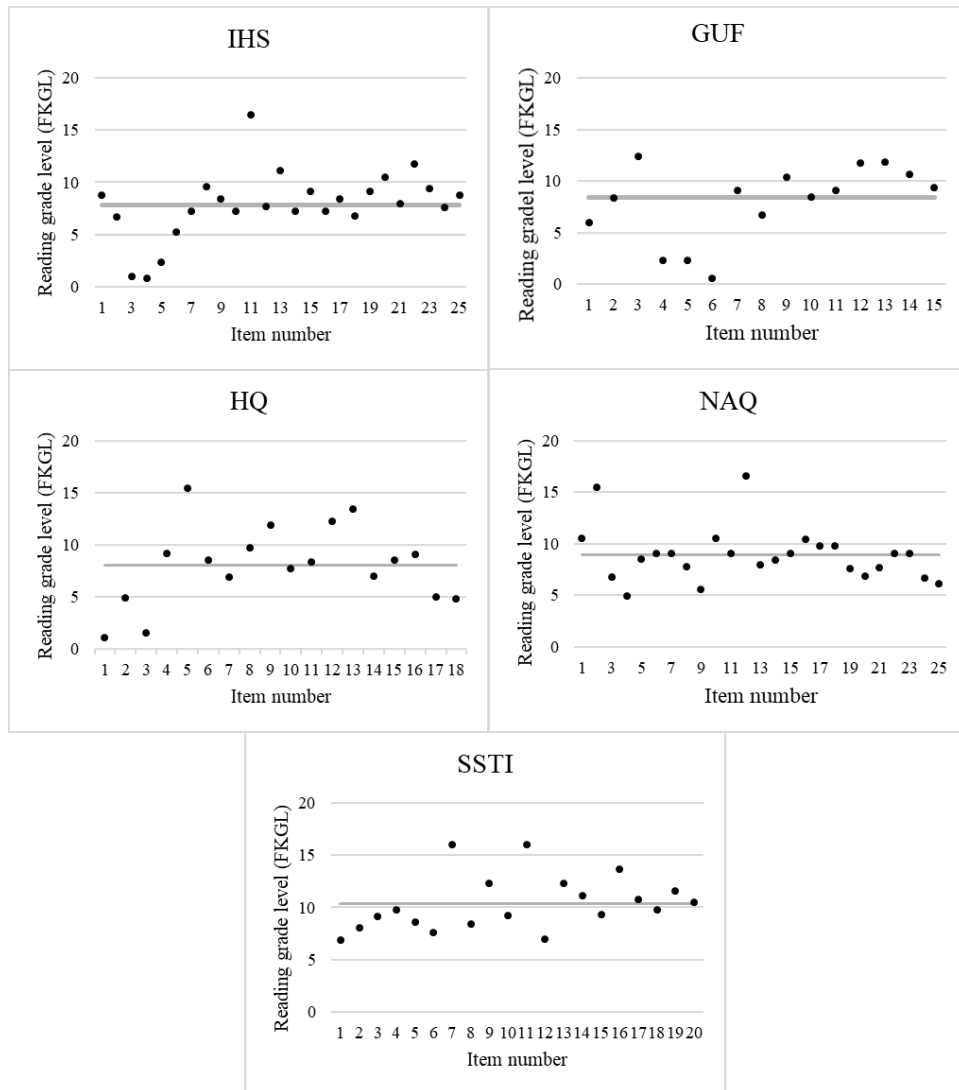


Figure 1. Readability grade levels for single items of questionnaires as per the FKGL formula. Solid light grey line represents the mean readability score for single items for each questionnaire. GUF- Geräuschüberempfindlichkeits-Fragebogen, NAQ- Noise Avoidance Questionnaire, HQ- Hyperacusis Questionnaire, SSTI- Sound Sensitive Tinnitus Index, IHS- Inventory of Hyperacusis Symptoms.