Research Article

Readability Level of Spanish-Language Patient-Reported Outcome Measures in Audiology and Otolaryngology

Laura Coco,^a Sonia Colina,^b Samuel R. Atcherson,^{c,d} and Nicole Marrone^a

Purpose: The purpose of this study was to examine the readability level of the Spanish versions of several audiology- and otolaryngology-related patient-reported outcome measures (PROMs) and include a readability analysis of 2 translation approaches when available—the published version and a “functionalist” version—using a team-based collaborative approach including community members.

Method: Readability levels were calculated using the Fry Graph adapted for Spanish, as well as the Fernandez-Huerta and the Spaulding formulae for several commonly used audiology- and otolaryngology-related PROMs.

Results: Readability calculations agreed with previous studies analyzing audiology-related PROMs in English

and demonstrated many Spanish-language PROMs were beyond the 5th grade reading level suggested for health-related materials written for the average population. In addition, the functionalist versions of the PROMs yielded lower grade-level (improved) readability levels than the published versions.

Conclusion: Our results suggest many of the Spanish-language PROMs evaluated here are beyond the recommended readability levels and may be influenced by the approach to translation. Moreover, improved readability may be possible using a functionalist approach to translation. Future analysis of the suitability of outcome measures and the quality of their translations should move beyond readability and include an evaluation of the individual's comprehension of the written text.

Health literacy refers to an individual's ability to understand and apply information provided in a health context, including, but not limited to, the ability to read (Kutner, Greenberg, & Baer, 2006; Ratzan & Parker, 2000). Examples include navigating a questionnaire, medication instructions, or comprehension of diagnosis (Institute of Medicine, 2004). Since the 1960s, research has demonstrated that low health literacy interferes with an individual's ability to manage diseases such as diabetes (Schillinger et al., 2002) and heart failure (Friel, 2016) and can predict negative surgical outcomes (Scarpato et al., 2016). Moreover, health literacy has been cited as the single greatest

predictor of an individual's health status (Baker et al., 2007), and it is an indicator of increased hospital visits, morbidity, mortality, and high health care costs (Baker et al., 2002; Institute of Medicine, 2004; Schillinger et al., 2002; Wolf, Gazmararian, & Baker, 2005).

Inadequate health literacy rates are a major public health challenge in the United States. A report from the National Assessment of Adult Literacy found health literacy had declined over the past 10 years, and 35% of American adults (85 million people) had basic or below basic levels of health literacy, meaning they had difficulty with tasks such as following instructions on a prescription and answering questions about a pamphlet they just read (Kutner, Greenberg, Jin, & Paulsen, 2006). It is estimated that approximately one third of Americans have limited health literacy proficiency that interferes with their ability to participate in the health care system (Parker, Ratzan, & Lurie, 2003). Improving the nation's health literacy rates is becoming an increasingly crucial issue in public policy and was made a priority of the United States' Healthy People 2010 national action plan (U.S. Department of Health and Human Services, 2010). Within audiology and otolaryngology, similar recommendations have been made on strengthening publicly available information on hearing and hearing health care across all

^aDepartment of Speech, Language, and Hearing Sciences, University of Arizona, Tucson

^bDepartment of Spanish and Portuguese, University of Arizona, Tucson

^cDepartment of Audiology and Speech Pathology, University of Arkansas at Little Rock

^dUniversity of Arkansas for Medical Sciences, Little Rock

Correspondence to Nicole Marrone: nmarrone@email.arizona.edu

Editor-in-Chief: Sumitrajit Dhar

Editor: Lauren Calandrucio

Received February 16, 2017

Revision received May 3, 2017

Accepted May 8, 2017

https://doi.org/10.1044/2017_AJA-17-0018

Disclosure: The authors have declared that no competing interests existed at the time of publication.

health literacy levels (National Academies of Science, Engineering, and Medicine, 2016). The American Academy of Audiology's *Standards of Practice* also recommends that audiologists use materials that are at appropriate health literacy levels (American Academy of Audiology, 2012). A recent cross-sectional study by Megwalu and Lee (2016) demonstrated that at least 10% of adult patients seen in a tertiary care otolaryngology clinic had low health literacy. Further, 26% of patients had difficulty with one of three health literacy domains assessed (filling out medical forms, reading medical material, or understanding written information about their medical condition).

In audiology and otolaryngology, self-assessment questionnaires and patient-reported outcome measures (PROMs) are commonly used in clinical research and in practice settings to guide aural rehabilitation, perform needs assessment, evaluate degree of disability, and determine the effectiveness of an intervention. It is unfortunate that research has shown these measures are not sensitive to the nation's health literacy levels. It is recommended that health-related materials for the general public should be written at or below a fifth-grade level, or age 10–11 years (Doak, Doak, & Root, 1996; Weiss & Coyne, 1997). However, a number of previous studies have shown that widely used English-language materials in audiology and otolaryngology are beyond the recommended level (Atcherson, Zraick, & Brasseux, 2011; Kahn & Pannbacker, 2000; Kelly-Campbell, Atcherson, Zimmerman, & Zraick, 2012; Laplante-Lévesque & Thorén, 2015; Nair & Cienkowski, 2010).

Further complicating this health literacy mismatch, there is evidence to suggest readability is affected as a result of translating a written text from one language to another (Ciobanu, Dinu, & Pepelea, 2015). Due to an increased prevalence of international and multilingual research, there is a crucial need to translate existing materials. In the United States, Spanish is overwhelmingly the leading non-English language. An estimated 37 million individuals over the age of 5 years speak Spanish at home (Rumbaut & Massey, 2013), and this number is expected to increase to up to 42 million by the year 2020 (Shin & Ortman, 2011). Therefore, Spanish-language materials, including PROMs, will be increasingly needed in health care settings, including audiology and otolaryngology. In an effort to serve a multilingual and diverse patient cohort, it is crucial to analyze whether Spanish-language PROMs are written at a level accessible to the general public. The present study is the first investigation of the readability level of several audiology- and otolaryngology-related PROMs that have been translated to Spanish.

Readability is a widely used mathematical calculation used to assess text complexity. Dozens of formulae exist to calculate readability in English, yet there is no standard method for choosing a formula (for a review of formulae in English, see Ley & Florio, 1996). Readability formulae for Spanish texts are typically adaptations of English formulae, and their availability is limited. In both languages, most formulae calculate readability on the basis of the same core

measures: average sentence length and number of words, and the assumption that a text is easier to read if the sentences and words are shorter. The resulting score is typically equated to a difficulty level or a school grade reading level on the basis of the United States education system, identifying the level someone would need to read the text. For this study, we used three standard Spanish formulae on the basis of their availability and prevalence in the literature: the Fernandez-Huerta Formula, the Fry Graph adapted for Spanish, and the Spaulding Formula. Variability in readability estimations is possible, and even likely, given that each formula gives weight to different semantic features. Therefore, it is recommended that more than one formula is used, and results are presented across formulae (Ley & Florio, 1996).

Reading Formulae in Spanish

Detailed explanation of each formula can be found in the Appendix. The Fernandez-Huerta Formula (Fernández Huerta, 1959) is a Spanish adaptation of the Flesch Reading Ease score for analyzing texts in English, one of the oldest and most common readability formulae (Flesch, 1948). The readability calculation is based on syllable count and sentence length in a text sample of the first 100 words. The output is an index from 0–100, with a lower number representing a more difficult text. Difficulty levels can then be converted to U.S. reading grade levels.

The Fry Graph calculation was initially created for English texts, but was later adapted for Spanish with the addition of a correction factor (Gilliam, Peña, & Mountain, 1980). Using this procedure, three 100-word passages are selected. The average number of syllables and average number of sentences per selection are calculated. These two values are plotted on a Standard Fry Graph, and their intersection determines the U.S. reading grade level appropriate for the text.

Last, the Spaulding Formula is based on word usage as measured by calculating word density, a percentage of words not appearing among the most frequently used words in Spanish (list provided in Fountain-Chambers, 1982), and sentence complexity as measured by the average sentence length in the sample. The output is a difficulty level index, for which corresponding categories of difficulty are provided. SPANREAD, a computer program used to calculate the Spaulding Formula, provides estimated corresponding U.S. reading grade levels using Thonis (1976) designations (Fountain, 1989).

Translation Approaches and Quality Assessment

Translation methodologies can vary, ranging from literal, one-step structural approaches to more dynamic, functional multistep processes. A commonly used tool for quality assurance in translation is back-translation, which involves translating the target text back to the primary language to verify its accuracy (Brislin, 1970, 1986; Tyupa, 2011). Published versions of PROMs in health-related research

typically use a combination of back-translation and other translation quality assessment methods (e.g., translation panels, pretesting). Colina, Marrone, Ingram, and Sánchez (2016) recently reviewed the limitations of back-translation as related to translation quality assessment. They point out that back-translation relies on an outdated, 1970s conceptualization of translation. Research in translation studies, starting at least in the 1980s, has shown that this quality assessment method relied on a restricted notion of translation that focused almost exclusively on the written text as an object, seen as a sequence of sentences, and on the linguistic structure of these sentences. Essential elements in translation and in text production and reception such as audience, purpose, social conditions of the text, or the translation were rarely, if ever, considered. Furthermore, back-translation is highly dependent on *equivalence*, a hard-to-define, controversial, and unstable construct (Colina, 2015; Herdman, Fox-Rushby, & Badia, 1997). Most notions of equivalence in health care translations are word- or conceptually based and do not contemplate the reader, the context of the translation, or the text as a unit (Colina, 2015).

Although back-translation can be useful to spot errors in the translation of purely referential meaning and one-to-one meaning correspondences, such as specialized technical terminology (e.g., translation of chemical compounds in a chemistry paper), it can actually work in the opposite fashion it was intended for other types of text. As Brislin (1970) noted, the more literal the translation, the closer the back-translation will be to the source. Therefore, back-translation tends to favor literal translations, identifying them as “better” translations. A literal translation can of course be highly problematic in reader-oriented texts, such as marketing or health promotion and questionnaires or outcome measures, whose main goal is to obtain the requested information from a person. In addition, a highly literal translation can pose severe readability and comprehension problems for a monolingual target language reader (Colina et al., 2016).

Functionalist Translation Approach

Given the limitations of back-translation and the equivalence approach it presupposes, Colina et al. (2016) demonstrated an innovative alternative method using a case study in an audiology research context—that is, a functionalist-collaborative approach. It is referred to as functionalist because it is based on the functionalist theory of translation (Nord, 1997) according to which translation decisions are guided by the purpose the translation is expected to achieve. Colina et al. (2016) proposed that each text should be translated according to the type of text involved and to the purpose that the research team wants to attain by means of the translation. In other words, the purpose that the translation is intended to achieve will determine the translational approach—that is, the degree of correspondence between the source text and the target text, as well as the translation decisions made by the translation team. In addition, the actual form of the translated text is determined by other contextual factors surrounding

the translation, such as the target audience and their characteristics, including socioeconomic and geographic considerations, the medium (e.g., Is the translation all in writing or is it to be read to the user by an interviewer?), and the purpose (e.g., to inform readers of a legal requirement, to educate them, to move them to action, etc.). This is in fact not different from what takes place in writing original texts for which authors are well aware of their purpose and their audience. The obvious difference is that, in translation, the content is provided as a source text that the translator/translation team is required to follow as long as the purpose and instructions for the translation allow it (Colina et al., 2016). The functionalist approach is in stark contrast to traditional approaches that aim for linguistic equivalence between the translation and the source, and often fail to take contextual and audience needs into account.

In Colina et al. (2016), an interdisciplinary team made translation decisions on the basis of the cultural and linguistic needs of the target population. Members of the translation team included researchers from the areas of public health and audiology, as well as social and health workers, translation studies experts, and community members with expertise on the text’s topic. In accordance with the functionalist methodology, the team translated the source text on the basis of factors, such as characteristics of the target audience (e.g., educational level and linguistic status), delivery medium (written or read aloud), and purpose of the source text (e.g., informational, educational, outcomes assessment). Additional details about the translation methodology can be found in Colina et al. (2016). Three PROMs translated from the Colina et al. (2016) process met our inclusion criteria for this study (minimum word count needed for a readability analysis), and were therefore included in our readability analysis.

Current Study

The purpose of this study was to examine the readability level of five PROMs used in audiology and otolaryngology and identify if they exceeded the fifth-grade reading level suggested for health-related materials written for the general public. Our analysis included a comparison of the readability level of several PROMs using two translation approaches. To be specific, the analysis included a version on the basis of the original published translation, which often reveals a more literal approach, perhaps with some adaptation of culturally marked terms and frequently in combination with back-translation as a quality control method, and one created using the collaborative functionalist approach described in Colina et al. (2016).

Method

This study utilized a nonexperimental, descriptive design. Readability calculations were considered for the following audiology- and otolaryngology-related PROMs in Spanish: International Outcomes Inventory–Hearing Aids

(IOI-HA) published version, International Outcomes Inventory–Alternative Interventions (IOI-AI) functionalist version, Hearing Handicap Inventory for the Elderly–Screening Version (HHIE-S) published and functionalist versions, the Abbreviated Profile of Hearing Aid Benefit (APHAB) published version, the Self-Efficacy for Situational Communication Management Questionnaire (SESMQ) functionalist version, and the Nijmegen Cochlear Implant Questionnaire (NCIQ) published version (refer to the Appendix for additional details and references for each PROM).

Outcome measures were chosen on the basis of their clinical relevance and use in clinical research contexts within the last 5 years, as well as structural components of the questionnaire; a minimum of 100 words was needed to perform a readability analysis. Functional translations used in this study were created previously by a team of researchers and community members for the purpose of a community-based intervention to expand access to hearing health care. Further description on the functionalist translation method involving collaboration between content experts and community members can be found above and in Colina et al. (2016). Detailed explanation of the content areas and response format of each questionnaire can be found in Table 1.

Readability of each PROM was calculated using the three formulae (Fry, Fernandez-Huerta, and Spaulding) and the calculations are described in detail in the Appendix. PROM items, including question stems and response scales, were typed into a word document, concatenated, and reviewed against the original text for accuracy. Scoring for the Fry Graph formula and the Fernandez-Huerta formula calculations were performed by hand independently by a bilingual audiologist (L.C.) and a bilingual audiology graduate student. Any discrepancies in the scoring were

reconciled by discussion between these examiners. Scoring for the Spaulding Formula was performed by SPANREAD, a computer program (Fountain, 1989).

Results

Table 2 provides a description of the readability results for all questionnaires corresponding to each formula along with mean, standard deviation, and ranges, referring to U.S. reading grade level. According to the Fry Graph calculation, four PROMs exceeded the fifth-grade reading level recommended for the general public: published IOI-HA (ninth), functionalist IOI-AI (eighth), published HHIE-S (sixth), and published APHAB (seventh). For the Spaulding calculation, one PROM exceeded this benchmark (published APHAB, sixth/eighth), yet five other PROM reading levels were measured at fifth grade. Last, according to results of the Fernandez-Huerta Index, all PROMs exceeded the recommended fifth-grade benchmark with the exception of the functionalist version of the SESMQ, which received a readability level of fifth grade. Averaging across the three readability formulae for each PROM, the published version of the IOI-HA was the most difficult text, with a readability level of eighth/ninth grade. The least challenging was the functionalist version of the SESMQ, with an average readability level of fourth grade.

Analysis Across Translation Versions

The IOI and the HHIE-S were the only PROMs with translations available in both published and functionalist versions. For both PROMs, the functionalist versions yielded lower (easier) readability grade levels than the published versions. Averaging the readability levels across all

Table 1. Patient-reported outcome measures (PROMs) with description and published translation method.

PROM	Description	Published translation method
International Outcomes Inventory–Alternative Interventions (Noble, 2002)	Seven-item questionnaire to assess real-world effects of auditory rehabilitation services using alternative interventions.	Translation and review (Cox et al., 2002)
Hearing Handicap Inventory for the Elderly–Screening Version (Lichtenstein et al., 1988)	10-item questionnaire evaluating the perceived emotional and social handicap of hearing loss.	Translation and back-translation (Lichtenstein & Hazuda, 1998)
Abbreviated Profile of Hearing Aid Benefit (Cox & Alexander, 1995)	24-item questionnaire comparing the patient's benefit with and without amplification categorized by four subscales: ease of communication, reverberation, background noise, and aversiveness.	Translation and back-translation (Hearing Aid Research Laboratories, n.d.)
Self-Efficacy for Situational Communication Management Questionnaire (Jennings et al., 2014)	20-item questionnaire measuring communication impact in adults with acquired hearing loss by their reported ability to hear and degree of self-efficacy in 20 situations.	N/A
Nijmegen Cochlear Implant Questionnaire (Hinderink et al., 2000)	60-item questionnaire of health-related assessment of quality of life changes experienced by cochlear implant users using three general domains: physical, psychological, and social.	Translation and back-translation (Sanchez-Cuadrado et al., 2015)

Table 2. Results of readability calculations.

PROM	Fry Graph	Spaulding	Fernandez-Huerta	<i>M</i>	<i>SD</i>	Range
Published translation versions						
IOI-HA	ninth	72.1 (fifth)	56.6 (10th/12th)	eighth/ninth	3.06	fifth to 12th
HHIE-S	sixth	77.3 (fifth)	59.1 (10th/12th)	seventh	3.21	fifth to 10th/12th
APHAB	seventh	83.7 (sixth/eighth)	70.0 (eighth/ninth)	eighth	0.87	sixth/eighth to sixth/ninth
NCIQ	fourth	76.0 (fifth)	81.9 (sixth)	fifth	1.00	fourth to sixth
Functionalist translation versions						
IOI-AI	eighth	60.0 (third)	44.3 (college)	eighth	5.00	third to college
HHIE-S	third/fourth	73.7 (fifth)	81.6 (sixth)	fifth	1.26	third/fourth to sixth
SESMQ	second	59.0 (third/fifth)	95.1 (fifth)	fourth	1.53	second to fifth

Note. PROM = patient-reported outcome measures; IOI-HA = International Outcomes Inventory for Hearing Aids; HHIE-S = Hearing Handicap Inventory for the Elderly–Screening version; APHAB = Abbreviated Profile of Hearing Aid Benefit; NCIQ = Nijmegen Cochlear Implant Questionnaire; IOI-AI = International Outcomes Inventory for Alternative Interventions; SESMQ = Self-Efficacy for Situational Communication Management Questionnaire.

three formulae for the HHIE-S published version produced a mean reading level of seventh grade, whereas the mean reading level for the functionalist version was fifth grade. The mean reading level for the IOI published version was eighth/ninth grade, whereas the mean reading level for the functionalist version was eighth grade. The third functionalist PROM included in the analysis, the SESMQ, yielded a mean reading level of fourth grade, but lacked a published Spanish translation for comparison.

Discussion

In the present study, we calculated the readability level for a sample of commonly used Spanish language audiology- and otolaryngology-related PROMs and discovered many were beyond the recommended fifth-grade level for health-related materials. In the United States, an estimated 21% of adults read below a fifth-grade level (Kutner, Greenberg, & Baer, 2006). In addition, national health literacy data suggest only half of adults are able to access, process, and apply health information provided to them in written form beyond a basic level (Kutner, Greenburg, Jin, & Paulsen, 2006). Taking these demographics into consideration, there are potential problems that could arise if a mismatch occurs between the patient's reading level and the reading level of the written materials administered to them in clinical and research settings. If such incompatibilities exist, there may be a misrepresentation of the patient's perspective due to inadequate comprehension, which has been documented in the literature to lead to mismanagement of disease (Williams, Baker, Parker, & Nurss, 1998) or misreporting on clinical tests (Weiss, Reed, & Kligman, 1995).

Our results are in line with previous studies that have calculated the readability index of English-language texts in audiology and otolaryngology and found most measures were beyond the level recommended for the average patient (Atcherson et al., 2011; Kahn & Pannbacker, 2000; Kelly-Campbell et al., 2012; Laplante-Lévesque & Thorén, 2015; Nair & Cienkowski, 2010). In one recent study, researchers examined self-reported hearing disability measures

in English and concluded, using three formulae, the range of readability for the HHIE was sixth to 9.6th grade and the range for the APHAB was sixth to 7.8th grade (Kelly-Campbell et al., 2012). Our results for the published Spanish translations of these measures were comparable; the range of readability levels for the HHIE-S was fifth to 10th/12th grade and the range for the APHAB was sixth/eighth to eighth/ninth grade.

Variability Across Formulae

Across all three formulae, most PROMs evaluated in this study were beyond the recommended fifth-grade reading level for health-related materials. However, the difficulty level of the text was dependent on the formula used. For example, the readability of IOI-AI (functionalist) ranged from third-grade to college-level difficulty using the Spaulding and Fernandez-Huerta formulae, respectively. Differences in readability scores between the formulae are to be expected, due to different factors included in each calculation, including word count, word length as estimated by syllable count, and sentence density, and the varying weight that is given to each factor. The Spaulding Formula is the only calculation based on word frequency, a factor determined in part by the percentage of words not appearing in a list of most frequently used words in Spanish. Many audiology- and otolaryngology-related terminology included in the questionnaires are considered "rare words" not appearing on this list, so PROMs with more challenging, obscure words, regardless of syllable count, would yield higher readability estimations on the basis of the Spaulding Formula compared with the Fry Graph and Fernandez-Huerta formulae. However, according to the Spaulding method, if a word is repeated four or more times within a text sample, the difficulty of the word is weighted lower, under the assumption the reader is becoming more familiar with the material with increased repetition and exposure (Spaulding, 1956). However, the audience may be unfamiliar with technical terminology regardless of the number of times the word is repeated, so a lower readability level might therefore be misleading in some cases with the

Spaulding method. Variation across formulae highlights the importance of cross-referencing the results of multiple formulae, thus producing a more holistic understanding of the text's complexity.

Readability Difficulty Across Translation Versions

An analysis of two translation approaches demonstrated PROMs translated using a functionalist methodology yielded generally lower (easier) readability levels than the published versions, implying the team-based translation approach may better meet the needs of populations at risk for low health literacy levels. According to readability calculations performed for 12 calculations for the published version PROMs, only one (the Nijmegen Cochlear Implant Questionnaire) was below the recommended reading level of below fifth-grade literacy level. According to readability levels across all three formulae, the SESMQ (functionalist) appears accessible for populations of low health literacy. In contrast, the IOI-HA (published) is perhaps inappropriate for use in groups that have low health literacy. Out of nine calculations performed for PROMS translated using a functionalist approach, four were below the fifth-grade literacy level. This suggests the functionalist approach helped affect readability by shortening sentences and words and increasing the number of frequently heard words. However, further research is needed to identify if the functionalist translation approach affects improved comprehension of the text.

It is important to recognize the strengths and limitations of the readability calculations. Readability is a popular and informative analysis because it is simple, objective, and has been shown to correlate to reading comprehension tasks (Fry, 1968). Useful applications include text simplification (Kushalnagar et al., 2016) and evaluating the suitability of a certain application of materials (Lau & King, 2006). However, because readability formulae rely solely on the structural properties of the text, calculations fail to reflect several factors that could affect the reader's comprehension (Friedman & Hoffman-Goetz, 2006). These factors might include: inference load, pragmatic, and socio- or psycholinguistic properties of the text, as well as the reader's attention level, attitude, and conceptual understanding of the material (Enkvist, 1991; Ferstl, Neumann, Bogler, & Von Cramon, 2008; Kemper, 1983). Furthermore, readability formulae were initially developed for analyzing large blocks of prose, and therefore require a minimum word count to perform calculations. This prevents calculating readability for questionnaires with limited text.

Although readability formulae for English texts are abundant, the availability for Spanish texts is limited. In addition, computer-based analysis is becoming increasingly prevalent, allowing for an efficient and accessible way to calculate readability, although the availability for Spanish texts is still limited. In the current study, we chose three readability formulae on the basis of their prevalence in the literature, and only one (Spaulding Formula) was available in a computer-based format. Manual calculations of readability can be time-consuming to complete and carry the

risk of human error. However, caution must also be applied when using computer-based formulae. It is crucial to understand the underlying factors weighed in the calculation to thoroughly understand the results of the analysis.

Despite its limitations, obtaining a readability calculation is a reasonable starting place when evaluating the overall suitability of PROMs and other health-related materials. In addition to research, PROMs are widely used clinically to help inform the effectiveness of a therapy or guide rehabilitation, for example. In addition, there is growing international interest regarding the use of PROMs in the electronic medical records system and in reimbursement models (Black, 2013; Centers for Medicare & Medicaid Services, 2016). In each of these arenas, it is crucial that materials are compatible with the linguistic needs of the intended audience so that the data collected are reliable and patient safety is upheld.

Going forward, it will be crucial to consider other aspects that contribute to comprehension of the questionnaire, especially when working with populations with low health literacy, or who may be unfamiliar with the research process or clinical assessment of outcomes. For example, it may be important to consider a questionnaire's response style. The commonly used Likert scale has shown to be problematic in cross-cultural research, possibly due to poor translation and differences in how groups tend to respond to scaled response items (D'Alonzo, 2011). In addition, social desirability bias, or the tendency to self-report positively, is a documented issue in cross-cultural research and should be considered (Keillor, Owens, & Pettijohn, 2001). Once a readability level is determined, additional steps may be taken to further evaluate the materials in domains such as construct validity and cultural appropriateness. Tools that are sensitive to these dimensions include the Suitability Assessment of Materials, a systematic guide to rate health-related materials on their appropriateness (Doak et al., 1996), or a team-based approach to adapting materials for a particular audience, as demonstrated in the case study on functionalist translation by Colina et al. (2016). In addition, D'Alonzo (2011) suggested several ways to adapt questionnaires to suit the needs of cross-cultural populations, including offering a limited number of multiple-choice response options and utilizing community workers as facilitators to minimize social desirability bias for questionnaires administered via interview.

Although results from the present study demonstrate that several PROMs in Spanish are beyond the literacy level of the average patient, this is not meant to discourage the use of outcome measures. New PROMs or adaptations of existing measures are certainly needed that will be more easily readable and understood by the majority of patients. Modifications can be made to accommodate reading levels of current materials, though their psychometric properties may then need to be reevaluated. Using less complicated vocabulary and shorter sentences and assisting individuals in reading items could help individuals to process and understand questions as well as provide valid responses. In addition, supplementing the text with

graphical illustrations has proven to be an effective way to reinforce key concepts and improve comprehension (Houts, Doak, Doak, & Loscalzo, 2006). However, such modifications may also potentially change responses and have implications on the resources needed to administer PROMs. In addition, more accessible PROMs could potentially contribute to more efficient care by promoting more accurate and time-saving assessments when individuals are better able to independently complete questionnaires.

As a first step, researchers and clinicians can consider measuring the health literacy level of a subset of the patient population to obtain information about the communication abilities of the individuals they are serving (Lee, Stucky, Lee, Rozier, & Bender, 2010; Megwalu & Lee, 2016). Next, they can review currently published PROMs and additional written materials to assess whether they are suitable for use in the intended population. This can be done using readability formulae and/or evaluation materials that include comprehension assessment. Materials can then be revised to reflect the population's health literacy demands. The result should be clearer, simple language structure that is appropriate for the population it is intended for. If necessary, the outcome scale will need to be adapted as well. Additional consideration should also be given to the cultural relevance of the measure's content. Colina et al. (2016) demonstrated that this form of language mediation is effective when it involves a collaborative team made up of members of the community, as well as content experts, and individuals from various linguistic backgrounds.

Evaluations and modifications can be done informally, though not without risk to generalizability; therefore, looking forward, this process could be systematized if networks of translation and quality assessment support were formally established. Such networks ideally would bring together teams to collaborate on making currently published PROMs culturally relevant and linguistically appropriate, and perhaps generating new PROMs, if needed. The need for updated measures will become increasingly crucial as cross-cultural research and service delivery to multilingual patient populations continue to grow more prevalent.

Acknowledgments

Research in this publication was supported by the National Institute on Deafness and Other Communication Disorders under Award Number R21/R33 DC013681 (PI: N.M.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors thank Paulina Bueno and Dr. Joanna Fountain for their assistance in scoring the readability calculations and Diane Cheek for her research assistance.

References

American Academy of Audiology. (2012). *Standards of practice for audiology*. Retrieved from <http://www.audiology.org/sites/default/files/documents/StandardsofPractice.pdf>

- Atcherson, S. R., Zraick, R. I., & Brasseur, R. E.** (2011). Readability of patient-reported outcome questionnaires for use with persons with tinnitus. *Ear and Hearing, 32*, 671–673.
- Baker, D. W., Gazmararian, J. A., Williams, M. V., Scott, T., Parker, R. M., Green, D., . . . Peel, J.** (2002). Functional health literacy and the risk of hospital admission among Medicare managed care enrollees. *American Journal of Public Health, 92*, 1278–1283.
- Baker, D. W., Wolf, M. S., Feinglass, J., Thompson, J. A., Gazmararian, J. A., & Huang, J.** (2007). Health literacy and mortality among elderly persons. *Archives of Internal Medicine, 167*, 1503–1509.
- Black, N.** (2013). Patient reported outcome measures could help transform healthcare. *BMJ, 346*, f167.
- Brislin, R. W.** (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1*, 185–216.
- Brislin, R. W.** (1986). The wording and translation of research instruments. In W. J. Lonner & J. W. Berry (Eds.), *Field methods in cross-cultural research* (pp. 137–164). Beverly Hills, CA: Sage.
- Centers for Medicare & Medicaid Services.** (2016). *CMS quality measure development plan: Supporting the transition to the Merit-Based Incentive Payment System (MIPS) and alternative payment models (APMs)*. Baltimore, MD: Author. Retrieved from <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/Final-MDP.pdf>
- Ciobanu, A. M., Dinu, L. P., & Pepelea, F.** (2015). Readability assessment of translated texts. In G. Angelova, K. Bontcheva, & R. Mitkov (Eds.), *Proceedings of Recent Advances in Natural Language Processing* (pp. 97–103). Shoumen, Bulgaria: INCOMA.
- Colina, S.** (2015). *Fundamentals of translation*. Cambridge, UK: Cambridge University Press.
- Colina, S., Marrone, N., Ingram, M., & Sánchez, D.** (2016). Translation quality assessment in health research: A functionalist alternative to back-translation. *Evaluation & the Health Professions, 40*(3), 267–293. <https://doi.org/10.1177/0163278716648191>
- Cox, R. M., & Alexander, G. C.** (1995). The abbreviated profile of hearing aid benefit. *Ear and Hearing, 16*, 176–186.
- Cox, R. M., Stephens, D., & Kramer, S. E.** (2002). Translations of the International Outcome Inventory for Hearing Aids (IOI-HA). *International Journal of Audiology, 41*, 3–26.
- D'Alonzo, K. T.** (2011). Evaluation and revision of questionnaires for use among low-literacy immigrant Latinos. *Revista Latino-Americana de Enfermagem, 19*, 1255–1264.
- Doak, C. C., Doak, L. G., & Root, J. H.** (1996). *Teaching patients with low literacy skills* (2nd ed.). Philadelphia, PA: Lippincott.
- Enkvist, N. E.** (1991). On the interpretability of texts in general and of literary texts in particular. In R. D. Sell (Ed.), *Literary pragmatics* (pp. 1–26). New York, NY: Routledge.
- Fernández Huerta, J.** (1959). Medidas sencillas de lecturabilidad. *Consigna, 214*, 29–32.
- Ferstl, E. C., Neumann, J., Bogler, C., & Von Cramon, D. Y.** (2008). The extended language network: A meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mapping, 29*, 581–593.
- Flesch, R.** (1948). A new readability yardstick. *Journal of Applied Psychology, 32*, 221–233.
- Fountain, J.** (1989). *SPANREAD* [Computer software]. Georgetown, TX: Bibliotechnics.
- Fountain-Chambers, J. F.** (1982). *Readability of primary-grade Spanish reading books: a correlational study of the Spaulding*

- Formula and the Fry Graph (adapted)*. (Unpublished doctoral dissertation, Texas Woman's University).
- Friedman, D. B., & Hoffman-Goetz, L.** (2006). A systematic review of readability and comprehension instruments used for print and web-based cancer information. *Health Education & Behavior, 33*, 352–373.
- Friel, C. J.** (2016). Improving health outcomes for low health literacy heart failure patients. *Home Healthcare Now, 34*, 434–439.
- Fry, E.** (1968). A readability formula that saves time. *Journal of Reading, 11*, 513–516, 575–578.
- Gilliam, B., Peña, S. C., & Mountain, L.** (1980). The Fry Graph applied to Spanish readability. *The Reading Teacher, 33*, 426–430.
- Hearing Aid Research Laboratory, University of Memphis** (n.d.). *APHAB – FORMULARIO A*. Retrieved from <http://harlmemphis.org/files/6113/4618/0943/SPANISH.pdf>
- Herdman, M., Fox-Rushby, J., & Badia, X.** (1997). 'Equivalence' and the translation and adaptation of health-related quality of life questionnaires. *Quality of Life Research, 6*, 237–247.
- Hinderink, J. B., Krabbe, P. F., & Van Den Broek, P.** (2000). Development and application of a health-related quality-of-life instrument for adults with cochlear implants: The Nijmegen cochlear implant questionnaire. *Otolaryngology—Head and Neck Surgery, 123*, 756–765.
- Houts, P. S., Doak, C. C., Doak, L. G., & Loscalzo, M. J.** (2006). The role of pictures in improving health communication: A review of research on attention, comprehension, recall, and adherence. *Patient Education and Counseling, 61*, 173–190.
- Institute of Medicine.** (2004). *Health literacy: A prescription to end confusion*. Washington, DC: National Academies Press.
- Jennings, M. B., Cheesman, M. F., & Laplante-Lévesque, A.** (2014). Psychometric properties of the self-efficacy for situational communication management questionnaire (SESMQ). *Ear and Hearing, 35*, 221–229.
- Kahn, A., & Pannbacker, M.** (2000). Readability of educational materials for clients with cleft lip/palate and their families. *American Journal of Speech-Language Pathology, 9*, 3–9.
- Keillor, B., Owens, D., & Pettijohn, C.** (2001). A cross-cultural/cross-national study of influencing factors and socially desirable response biases. *International Journal of Market Research, 43*, 63–84.
- Kelly-Campbell, R. J., Atcherson, S. R., Zimmerman, K. R., & Zraick, R. I.** (2012). Readability of audiologic self-report assessment tools. *Journal of the Academy of Rehabilitative Audiology, 45*, 63–73.
- Kemper, S.** (1983). Measuring the inference load of a text. *Journal of Educational Psychology, 75*, 391–401.
- Kushalnagar, P., Smith, S., Hopper, M., Ryan, C., Rinkevich, M., & Kushalnagar, R.** (2016). Making cancer health text on the Internet easier to read for deaf people who use American Sign Language. *Journal of Cancer Education*. Advance online publication. <https://doi.org/10.1007/s13187-016-1059-5>
- Kutner, M., Greenberg, E., & Baer, J.** (2006). *A first look at the literacy of America's adults in the 21st century*. (NCES 2006-470). Washington, DC: National Center for Education Statistics. Retrieved from <https://nces.ed.gov/naal/pdf/2006470.pdf>
- Kutner, M., Greenburg, E., Jin, Y., & Paulsen, C.** (2006). *The health literacy of America's adults: Results from the 2003 National Assessment of Adult Literacy*. (NCES 2006-483). Washington, DC: National Center for Education Statistics. Retrieved from <https://nces.ed.gov/pubs2006/2006483.pdf>
- Laplante-Lévesque, A., & Thorén, E. S.** (2015). Readability of Internet information on hearing: Systematic literature review. *American Journal of Audiology, 24*, 284–288.
- Lau, T. P., & King, I.** (2006). Bilingual web page and site readability assessment. In *Proceedings of the 15th International Conference on World Wide Web* (pp. 993–994). New York, NY: ACM.
- Lee, S. Y. D., Stucky, B. D., Lee, J. Y., Rozier, R. G., & Bender, D. E.** (2010). Short Assessment of Health Literacy—Spanish and English: A comparable test of health literacy for Spanish and English speakers. *Health Services Research, 45*, 1105–1120.
- Ley, P., & Florio, T.** (1996). The use of readability formulas in health care. *Psychology, Health & Medicine, 1*, 7–28.
- Lichtenstein, M. J., Bess, F. H., & Logan, S. A.** (1988). Diagnostic performance of the hearing handicap inventory for the elderly (screening version) against differing definitions of hearing loss. *Ear and Hearing, 9*, 208–211.
- Lichtenstein, M. J., & Hazuda, H. P.** (1998). Cross-cultural adaptation of the Hearing Handicap Inventory for the Elderly—Screening Version (HHIE-S) for use with Spanish-speaking Mexican Americans. *Journal of the American Geriatrics Society, 46*, 492–498.
- Megwalu, U. C., & Lee, J. Y.** (2016). Health literacy assessment in an otolaryngology clinic population. *Otolaryngology—Head and Neck Surgery, 155*, 969–973.
- Nair, E. L., & Cienkowski, K. M.** (2010). The impact of health literacy on patient understanding of counseling and education materials. *International Journal of Audiology, 49*, 71–75.
- National Academies of Sciences, Engineering, and Medicine.** (2016). *Hearing health care for adults: Priorities for improving access and affordability*. Washington, DC: National Academies Press.
- Noble, W.** (2002). Extending the IOI to significant others and to non-hearing-aid-based interventions. *International Journal of Audiology, 41*, 27–29.
- Nord, C.** (1997). *Translating as a purposeful activity: Functionalist approaches explained*. Manchester, UK: St. Jerome Pub.
- Parker, R. M., Ratzan, S. C., & Lurie, N.** (2003). Health literacy: A policy challenge for advancing high-quality health care. *Health Affairs, 22*, 147–153.
- Ratzan, S. C., & Parker, R. M.** (2000). Introduction. In C. R. Selden, M. Zorn, S. C. Ratzan, & R. M. Parker (Eds.), *National Library of Medicine current bibliographies in medicine: Health literacy* (NLM Pub. No. CBM 2000-1). Bethesda, MD: National Institutes of Health, U.S. Department of Health and Human Services.
- Rumbaut, R. G., & Massey, D. S.** (2013). Immigration & language diversity in the United States. *Daedalus, 142*, 141–154.
- Sanchez-Cuadrado, I., Gavilan, J., Perez-Mora, R., Muñoz, E., & Lassaletta, L.** (2015). Reliability and validity of the Nijmegen Cochlear Implant Questionnaire in Spanish. *European Archives of Oto-Rhino-Laryngology, 272*, 1621–1625.
- Scarpato, K. R., Kappa, S. F., Goggins, K. M., Chang, S. S., Smith, J. A., Jr., Clark, P. E., ... Moses, K. A.** (2016). The impact of health literacy on surgical outcomes following radical cystectomy. *Journal of Health Communication, 21*(Suppl. 2), 99–104.
- Schillinger, D., Grumbach, K., Piette, J., Wang, F., Osmond, D., Daher, C., ... Bindman, A. B.** (2002). Association of health literacy with diabetes outcomes. *Journal of the American Medical Association, 288*, 475–482.
- Shin, H. B., & Ortman, J. M.** (2011). *Language projections: 2010 to 2020*. Paper presented at the Federal Forecasters Conference, Washington, DC. Retrieved from https://www.census.gov/hhes/socdemo/language/data/acs/Shin_Ortman_FFC2011_paper.pdf
- Spaulding, S.** (1956). A Spanish readability formula. *The Modern Language Journal, 40*, 433–441.
- Thonis, E. W.** (1976). *Literacy for America's Spanish speaking children*. Newark, NJ: International Reading Association.

- Tyupa, S.** (2011). A theoretical framework for back-translation as a quality assessment tool. *New Voices in Translation Studies*, 7, 35–46.
- U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion.** (2010). *National Action Plan to Improve Health Literacy*. Washington, DC: Author. Retrieved from <https://health.gov/communication/initiatives/health-literacy-action-plan.asp>
- Weiss, B. D., & Coyne, C.** (1997). Communicating with patients who cannot read. *The New England Journal of Medicine*, 337, 272–274.
- Weiss, B. D., Reed, R. L., & Kligman, E. W.** (1995). Literacy skills and communication methods of low-income older persons. *Patient Education and Counseling*, 25(2), 109–119.
- Williams, M. V., Baker, D. W., Parker, R. M., & Nurss, J. R.** (1998). Relationship of functional health literacy to patients' knowledge of their chronic disease: A study of patients with hypertension and diabetes. *Archives of Internal Medicine*, 158(2), 166–172.
- Wolf, M. S., Gazmararian, J. A., & Baker, D. W.** (2005). Health literacy and functional health status among older adults. *Archives of Internal Medicine*, 165, 1946–1952.

Appendix

Standard Formulas for Calculating Readability for Spanish Texts

Test	Formula
Fernandez-Huerta	Difficulty = $206.84 - (0.6 \times \text{Total Number of Syllables}) - 1.02 \times (\text{Total Number of Words})$
Spaulding	Difficulty = $1.609 (\text{Average Sentence Length}^a) + 331.8 (\text{Density}^b) + 22.0$
Fry Graph	Select three samples of 100-word passages. Take the average number of words per 100 words (x-axis) and average number of syllables per 100 words – 67 (y-axis) plotted onto a Standard Fry Graph.

^aAverage Sentence Length = average number of words per sentence in entire passage. ^bDensity = decimal percentage of words in the passage not appearing among the most frequently used words in Spanish provided in Fountain-Chambers, 1982.