



EDUCATION FOR HEALTH

BRIEF COMMUNICATION

Validation of the Translation of an Instrument to Measure Reliability of Written Information on Treatment Choices: A Study on Attention Deficit/Hyperactivity Disorder (ADHD)

A Montoya¹, N Llopis², I Gilaberte¹

¹Lilly Research Laboratories, Alcobendas, Spain

²Medical Science Liaison, TEVA PHARMA SLU, Alcobendas, Spain

Published: December 2011

Montoya A, Llopis N, Gilaberte I

Validation of the Translation of an Instrument to Measure Reliability of Written Information on Treatment Choices: A Study on Attention Deficit/Hyperactivity Disorder (ADHD)

Education for Health, Volume 24, Issue 3, 2011

Available from: <http://www.educationforhealth.net/>

ABSTRACT

Introduction: DISCERN is an instrument designed to help patients assess the reliability of written information on treatment choices. Originally created in English, there is no validated Spanish version of this instrument. This study seeks to validate the Spanish translation of the DISCERN instrument used as a primary measure on a multicenter study aimed to assess the reliability of web-based information on treatment choices for attention deficit/hyperactivity disorder (ADHD).

Methods: We used a modified version of a method for validating translated instruments in which the original source-language version is formally compared with the back-translated source-language version. Each item was ranked in terms of comparability of language, similarity of interpretability, and degree of understandability. Responses used Likert scales ranging from 1 to 7, where 1 indicates the best interpretability, language and understandability, and 7 indicates the worst. Assessments were performed by 20 raters fluent in the source language.

Results: The Spanish translation of DISCERN, based on ratings of comparability, interpretability and degree of understandability (mean score (SD): 1.8 (1.1), 1.4 (0.9) and 1.6 (1.1), respectively), was considered extremely comparable. All items received a score of less than three, therefore no further revision of the translation was needed.



Conclusion: The validation process showed that the quality of DISCERN translation was high, validating the comparable language of the tool translated on assessing written information on treatment choices for ADHD.

Keywords: Validation, translation, DISCERN, ADHD.

Introduction

Attention deficit/hyperactivity disorder (ADHD) is one of the most common chronic conditions of childhood characterized by developmentally inappropriate symptoms of inattention, hyperactivity and impulsivity. ADHD usually has a childhood onset of symptoms that typically results in a chronic and pervasive pattern of impairment in school, work, social and daily adaptive functioning¹. Medication is often the first intervention used to treat ADHD, but long-term adherence to pharmacologic treatment for ADHD is frequently poor^{2,3}, which is associated with poor symptom control and continued impairments in many domains in their lives.

Given that parents of patients with ADHD share decision-making regarding the treatments their children receive, it is important that parents can discriminate between reliable and unreliable sources of information about potential treatments. There are few standardized instruments for evaluating the quality of sources of information regarding treatment options, and even fewer instruments have been translated and validated in languages other than English for use in cross-cultural studies^{4,7}.

Cross-cultural research has specific methodological problems, most relating to translation quality and the comparability of results in different cultural and ethnic groups. In many cases, once the translation process is complete, the implementation of the new version of the instrument follows immediately. However, there is still an important stage to be carried out, the validation of the translation. The validation of the translation may be arduous and requires time and money. However, unless this process is successfully implemented, the validity of the research results where such a translated version is implemented may be unreliable^{8,9}.

DISCERN is an instrument designed to help patients assess the reliability of written information on treatment choices¹⁰ that has attracted considerable interest since its conception. This instrument was originally created in the English language. During its development and also in a number of subsequent studies, DISCERN has shown satisfactory psychometric properties when used by health professionals and patients¹¹⁻¹⁴. DISCERN consists of 16 questions. The first section (questions one–eight) evaluates the reliability of the information (e.g., “Is it clear what sources of information were used to compile the publication?”) and the second section (questions nine–15) considers the quality of the information on treatment choices (e.g., “Does it describe the benefits of each treatment?”). Five-point Likert scales ranging from 1 (no) to 5 (yes) accompany these items. The final section (question 16) assesses the overall rating of the publication on a five-point Likert scale ranging from 1 (low quality with serious or extensive shortcomings) to 5 (high quality with minimal shortcomings).

DISCERN has been selected as a primary measure on a multicenter study aimed to assess the reliability of web-based information on treatment choices for ADHD in Spain. As an early stage in that research, we translated the DISCERN instrument and we formally validated the translation into the Spanish language.



The purpose of this report is to outline the methodology of translation and validation of the translated version of this instrument. The psychometric analysis of DISCERN in its new versions is beyond the scope of this report.

Objective

To translate and validate the translation into the Spanish language of the DISCERN instrument, used as a primary measure on a multicenter study aimed to assesses the reliability of web-based information on treatment choices for ADHD.

Materials and Methods

Translation

For the translation of the DISCERN instrument we used the back-translation method. The DISCERN instrument was initially translated from English into Spanish by a native Spanish-speaking translator. The Spanish version was then back-translated into English by an independent native English-speaking translator who was blinded to the original version. All inconsistencies between the resulting English version and the original version were examined and the Spanish version was adjusted appropriately.

Validation of the translation

For the validation of the resulting Spanish translated version we used a modified version of the translation validation method proposed by Sperber et al⁸. for validating translated instruments, which has been widely used to validate translated questionnaires (for medical disorders, health-related quality of life, etc.).

This approach includes a step in the translation validation process in which the original source-language version is formally compared with the back-translated source-language version in terms of comparability of language and similarity of interpretability. We additionally introduced a third measure, the degree of understandability¹⁵. Figure 1 shows a flow diagram of the process.

Testing comparability, interpretability and the degree of understandability

Each item in the two versions (source-language and back-translated) of the instrument was ranked in terms of following three dimensions:

1. *Comparability of language*, which refers to the formal similarity of words, phrases and sentences.
2. *Similarity of interpretability*, which refers to the degree to which the two versions engender the same response even if the wording is not the same,
3. *Degree of understandability*, which refers to the degree of comprehension of the two versions even if the wording is different.

Likert scales ranging from 1 (extremely comparable/ similar/ understandable), to 4 (moderately comparable/ similar/ understandable), to 7 (not at all comparable/similar/understandable), were used for each dimension.

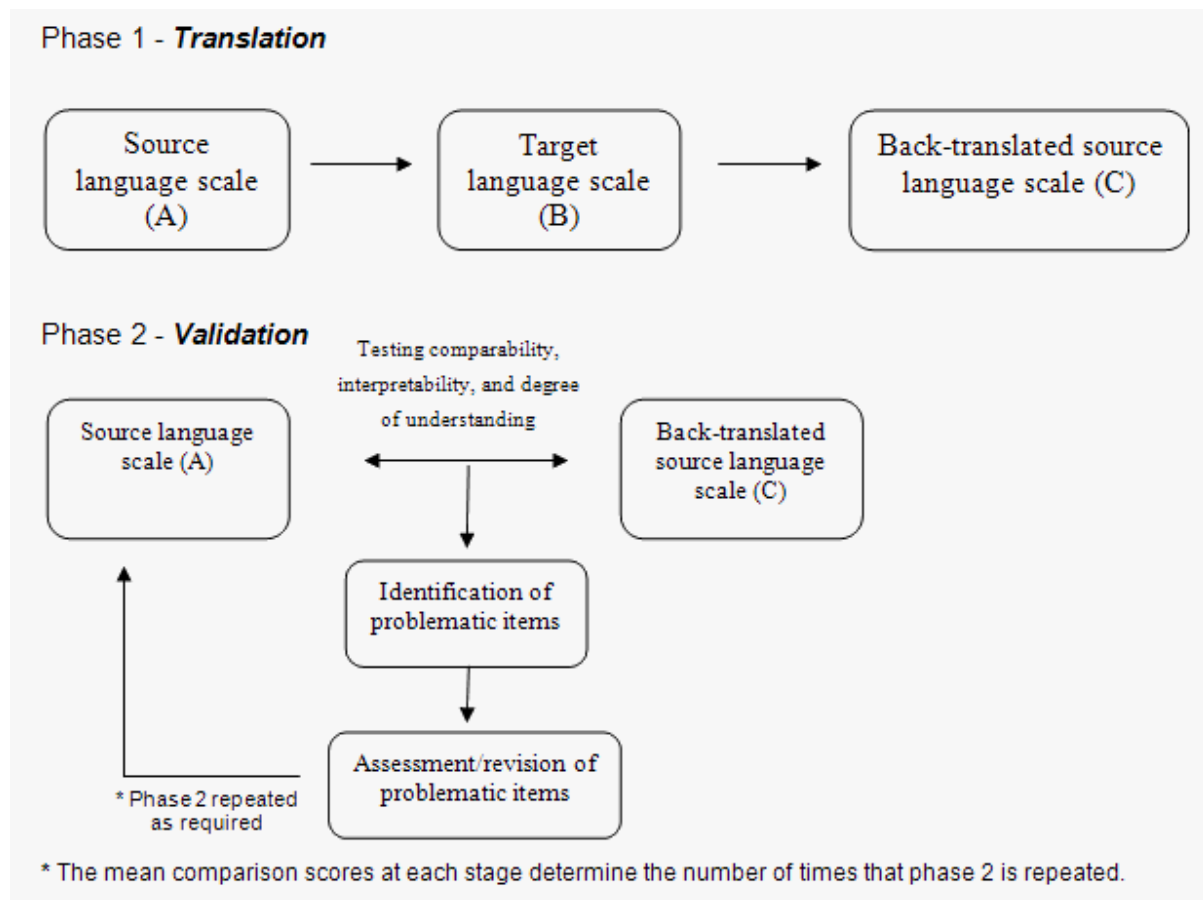


Figure 1. Flow diagram of the translation (phase1) and validation (phase2) processes

The two source-language versions are then compared. Any mean score greater than 3 requires a formal review of the translation. Any mean score between 2 and 3 is also considered problematic and has to be reviewed for possible correction. All problematic items are then retranslated until the item is comparable, interpreted, and understandable in the same manner in both languages. Figure 2 shows the rating scales used for this evaluation.

The ranking was performed by 20 raters with a mean age of 24 years who are, natives of Spain with a first language of Spanish and who learned English as a second language in college. All raters were postgraduate students in biomedical sciences who had to pass a test of English as a foreign language, required by their postgraduate school. They were independent of the investigators, and none were the translators of this instrument. This study was approved by an institutional review board.

Results

The mean score for each item pair (original and back-translated versions) in each of the three dimensions performed by the 20 raters are summarized in Table 1.



Please indicate the response which most closely represents how you would rate the following pairs of items in terms of:

Comparability of language	Extremely comparable			Moderately comparable			Not at all comparable
<i>How comparable is the formal wording?</i>	1	2	3	4	5	6	7
Similarity of Interpretation	Extremely interpretable			Moderately interpretable			Not at all interpretable
<i>Would the paired items be interpreted similarly, even if the wording is different?</i>	1	2	3	4	5	6	7
Degree of understandability	Extremely understandable			Moderately understandable			Not at all understandable
<i>Would the paired items be understood similarly?</i>	1	2	3	4	5	6	7

Please indicate only one response for each term and for each pair of items.

Figure 2. Comparability/Interpretability/Understandability rating sheet

None of the 16 items of DISCERN received a mean score of more than 3. Mean scores ranged from 1.1 to 2.85. The mean score and standard deviation for comparability, interpretability, and degree of understandability were 1.8 (1.1), 1.4 (0.9) and 1.6 (1.1), respectively. Figure 3 shows the mean score of the items for each term.

For some items mean scores fell between 2 and 3; these items were considered problematic and required review for possible correction. For example, item 2 in the original version is worded 'Does it achieve its aims?' while the back-translated version states that 'Are the stated objectives met?' These are clearly not the same even though the mean scores for comparability, similarity and degree of understandability were 2.3, 1.6 and 1.6 respectively. On reevaluation, it was found that the Spanish version was true to the original English and the problem lay in the back-translation, so the Spanish version was left unchanged.

The item 'Is it clear what sources of information were used to compile the publication (other than the author or producer)?' was translated to 'Are the sources for publication clearly stated (beyond the author and sponsor sources)?' The mean score for comparability of language was 2.4, showing that the raters believed there was a little discrepancy in language, as indeed there was. However, the mean score of similarity of interpretability was 2, indicating that despite the difference in formal language the reviewers did not believe there was a large interpretation problem. The mean score for degree of understandability was 2, indicating that although the wording was different, it was equally understood. On reviewing the Spanish version we observed that there was no serious problem with the translation so we did not change anything.

Item 5, 'Is it clear when the information used or reported in the publication was produced?' was translated as 'Is it clear when the referenced material is cited or used in the publication were produced?' and it received the highest (most problematic) scores. The mean score for degree of understandability was 2.9, the mean score for comparability of language was 2.7 and the mean score for similarity was 2.4, then the item was retranslated until this was considered as being interpreted, understandable and comparable in the same manner in both languages.



Table 1. Original and back-translated DISCERN items with mean comparison score for each item pair

Original English version	Back-translated English version	Comparability of language mean score (SD)	Similarity of interpretability mean score (SD)	Degree of understandability mean score (SD)
Are the aims clear?	Are the objectives clear?	2 (1.0)	1.8 (1.4)	1.9 (1.5)
Does it achieve its aims?	Are the stated objectives met?	2.3 (1.5)	1.6 (0.9)	1.6 (1.0)
Is it relevant?	Is it relevant?	1.6 (0.9)	1.3 (0.5)	1.3 (0.7)
Is it clear what sources of information were used to compile the publication (other than the author or producer)?	Are the sources for publication clearly stated (beyond the author and sponsor sources)?	2.4 (1.5)	2 (1.5)	2 (1.5)
Is it clear when the information used or reported in the publication was produced?	Is it clear when the referenced material is cited or used in the publication were produced?	2.7 (1.6)	2.4 (1.2)	2.9 (1.7)
Is it balanced and unbiased?	Is the publication balanced and unbiased?	1.7 (1.0)	1.3 (0.5)	1.8 (1.5)
Does it provide details of additional sources of support and information?	Does it provide details of additional sources of information or support?	1.6 (0.8)	1.2 (0.5)	1.3 (0.7)
Does it refer to areas of uncertainty?	Does the publication mention areas of uncertainty?	1.7 (0.7)	1.3 (0.5)	1.5 (0.8)
Does it describe how each treatment works?	Does it describe how each treatment option works?	1.5 (0.8)	1.2 (0.4)	1.3 (0.6)
Does it describe the benefits of each treatment?	Does it describe the benefits of each treatment?	1.5 (0.8)	1.2 (0.5)	1.4 (0.8)
Does it describe the risks of each treatment?	Does it describe the risk of each treatment?	1.2 (0.5)	1.1 (0.3)	1.2 (0.4)
Does it describe what would happen if no treatment is used?	Does it discuss the implications of not using a treatment?	1.8 (0.9)	1.4 (0.6)	1.5 (0.9)
Does it describe how the treatment choices affect overall quality of life?	Does it discuss how the treatment options affect quality of life in general?	1.8 (0.8)	1.5 (0.7)	1.3 (0.4)
Is it clear that there may be more than one possible treatment choice?	Is it clear that may be more than one treatment choice?	2 (1.5)	1.7 (1.1)	1.9 (1.5)
Does it provide support for shared decision-making?	The publication provides support for shared decision-making?	1.5 (0.7)	1.2 (0.5)	1.3 (0.7)
Based on the answer to all of the above questions, rate the overall quality of the publication as a source of information about treatment choices	Based on the answers to the previous questions, rate the overall quality of this publication as a source of information for treatment options	1.6 (0.9)	1.3 (0.9)	1.3 (0.9)

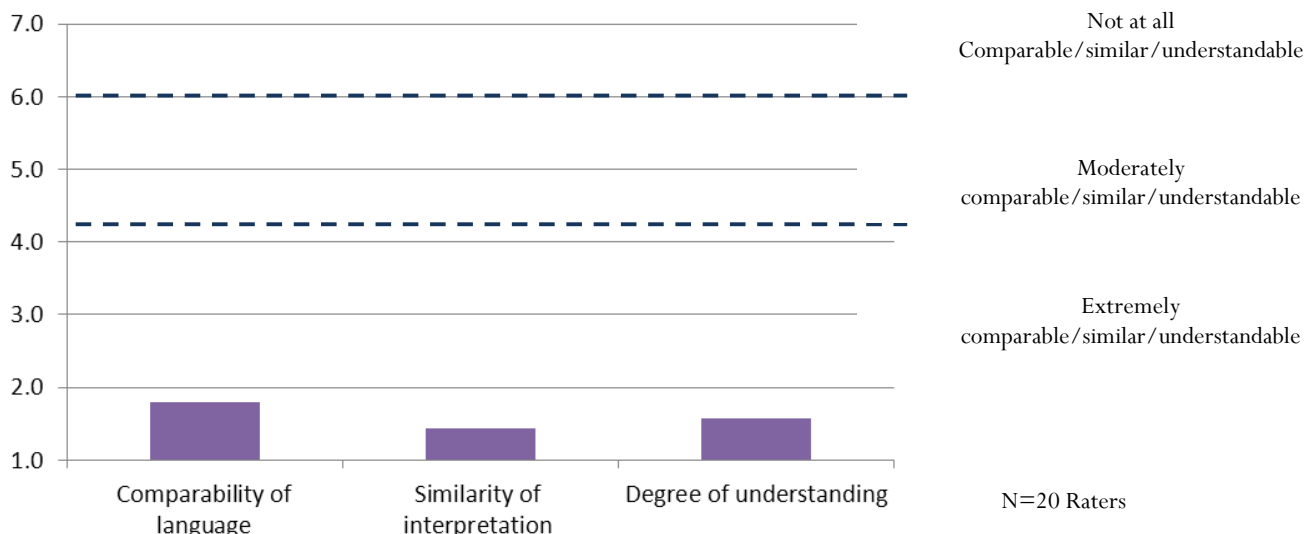


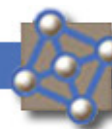
Figure 3. Total mean scores for comparability of language, similarity of interpretation and degree of understandability measurements

Discussion

There is a need for instruments to help patients and caregivers rate the reliability and usefulness of the information they find on the Internet and from other sources, with the ultimate aim of patients locating and using good quality information.. We translated and validated the translation into the Spanish language of the DISCERN instrument used as a primary measure on a multicenter study aimed to assess the reliability of web-based information on treatment choices for ADHD. Our results showed that the quality of the DISCERN translation was high, validating the comparable language of the tool translated into Spanish. .

Although there are some standardized instruments for evaluating the quality of sources of information regarding treatment options, most of them have been created in the English language and only a few have been translated in other languages. These instruments have received much criticism, mainly because they have not been properly validated, not only in the psychometric properties of the resulting translated version, but even earlier, during the translation stage, in the validation of the translation^{16,17}.

There are different approaches to the translation process. In the simplest method, a questionnaire is translated—often by unqualified translators—and the translated version is used without further validation. Another approach involves translation by committee. In this case, two or more translators work to produce a consensus questionnaire. Another approach is the back-translation method. In this case, a questionnaire is translated into the target language by one translator and then translated back into the source language by an independent translator who is blinded to the original questionnaire⁷. In this study we used the back-translation method. However, it is not enough to translate a questionnaire literally, even when using a back-translation process. One of the additional challenges is to validate the resulting translation in order to better adapt the translation in a culturally relevant and comprehensible form while maintaining the meaning and intent of the original items.



According to our data the DISCERN translation was considered extremely comparable. The mean score (s.d.) for comparability of language was 1.8 (1.1), indicating that the reviewers believed there was no major discrepancy in language, and despite the small differences in formal language there was no serious interpretation problem. The mean score (s.d.) for similarity of interpretation was 1.4 (0.9), indicating that the raters find similarity of interpretation even de wording was different. The mean score (s.d.) for the degree of understandability was 1.6 (1.1), showing that the raters found a high degree of understandability in each pair of items. As no ratings of the 16 items were qualified with a score of more than 3, no formal revision of the translation was required.

One of the limitations of this study is its relatively small sample size. Another is that although we ensured raters' proficiency in English they were not native English speakers. Their assessments of the comparability/understandability/similarity of the instrument in a language that is not their native tongue may not have been the same as assessments of native speakers.

The validation process showed that the quality of the DISCERN translation was high, supporting the validity and reliability of DISCERN as a tool for the assessment of written information on treatment choices for ADHD in the Spanish language. The validation of the translation can be arduous and requires investment of time and money, however, unless this process is successfully implemented, the validity of the research results where such a translated version is implemented may be unreliable.

Acknowledgement

We would like to thank Dr. Sasha Shepperd from the Research Centre, Department of Public Health, University of Oxford, Oxford for her kind feedback to the research and manuscript.

References

1. Biederman J, Faraone SV. Attention-deficit hyperactivity disorder. *Lancet*. 2005; 366(9481):237-248.
2. Spencer TJ, Biederman J, Mick E. Attention-deficit/hyperactivity disorder: diagnosis, lifespan, comorbidities, and neurobiology. *Ambulatory Pediatrics*. 2007; 7(1 Suppl):73-81.
3. Hugtenburg JG, Witte I, Heerdink ER. Determinants of compliance with methylphenidate therapy in children. *Acta Paediatrica*. 2006; 95(12):1674-1676.
4. Shumaker SA, Wyman JF, Uebersax JS, McClish D, Fantl JA. Health-related quality of life measures for women with urinary incontinence: the Incontinence Impact Questionnaire and the Urogenital Distress Inventory. Continence Program in Women (CPW) Research Group. *Quality of Life Research*. 1994; 3(5):291-306.
5. Vinsnes AG, Haltbakk J, Hunskaar S. A Norwegian version of the Incontinence Stress Questionnaire-Staff Reaction: translation and validation for cross-cultural use. *Journal of Nursing Measurement*. 2000; 8(1):71-86.
6. Badia Llach X, Castro Díaz D, Conejero Sagrañes J. [Validity of the King's Health questionnaire in the assessment of quality of life of patients with urinary incontinence. The King's Group]. *Medicina Clinica (Barc)*. 2000; 114(17):647-652.



7. Sperber AD. Translation and validation of study instruments for cross-cultural research. *Gastroenterology*. 2004; 126(1 Suppl 1):S124-S128.
 8. Cavusgil ST, Das A. Methodological Issues in Empirical Cross-Cultural Research: A Survey of the Management Literature and a Framework. : *Management International Review*. 1997; 37(1):71-96.
 9. Patrick DL, Martin ML, Bushnell DM, Marquis P, Andrejasich CM, Buesching DP. Cultural adaptation of a quality-of-life measure for urinary incontinence. *European Urology*. 1999; 36(5):427-435.
 10. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *Journal of Epidemiology and Community Health*. 1999; 53(2):105-111.
 11. Godolphin W, Towle A, McKendry R. Evaluation of the quality of patient information to support informed shared decision-making. *Health expectations*. 2001; 4(4):235-242.
 12. Rees CE, Ford JE, Sheard CE. Evaluating the reliability of DISCERN: a tool for assessing the quality of written patient information on treatment choices. *Patient Education and Counseling*. 2002; 47(3):273-275.
 13. Ademiluyi G, Rees CE, Sheard CE. Evaluating the reliability and validity of three tools to assess the quality of health information on the Internet. *Patient Education and Counseling*. 2003; 50(2):151-155.
 14. Griffiths KM, Christensen H. The quality and accessibility of Australian depression sites on the World Wide Web. *The Medical Journal of Australia*. 2002; 20(176):S97-S104.
 15. García-Borreguero D, Lahuerta-Dal Ré J, Albares J, Zaragoza S, de Gracia M. [Validated translation to Spanish of the evaluation questionnaire of patients with restless legs syndrome]. *Neurologia*. 2009; 24(10):823-834.
 16. Jadad AR, Gagliardi A. Rating health information on the Internet: navigating to knowledge or to Babel? *Journal of the American Medical Association*. 1998; 279(8):611-614.
 17. Gagliardi A, Jadad AR. Examination of instruments used to rate quality of health information on the internet: chronicle of a voyage with an unclear destination. *British Medical Journal*. 2002; 324(7337):569-573.
-