Linguistic Validation of the US Spanish Work Productivity and Activity Impairment Questionnaire, General Health Version

Mary C. Gawlicki, MBA,[1] Margaret C. Reilly, MPH,[2] Ana Popielnicki, BA,[1] Kate Reilly, MA[2]

[1]Corporate Translations, Inc., Mansfield Center, CT, USA; [2]Margaret Reilly Associates, Inc., New York, NY, USA

ABSTRACT

Introduction: There are no measures of health-related absenteeism and presenteeism validated for use in the large and increasing US Spanish-speaking population. Before using a Spanish translation of an available English-language questionnaire, the linguistic validity of the Spanish version must be established to ensure its conceptual equivalence to the original and its cultural appropriateness.
Objective: The objective of this study was to evaluate the linguistic validity of the US Spanish version of the Work Productivity and Activity Impairment questionnaire, General Health Version (WPAI:GH).
Methods: A US Spanish translation of the US English WPAI:GH was created through a reiterative process of creating harmonized forward and back translations by independent translators. Spanish-speaking and English-speaking subjects residing in the US self-administered the WPAI:GH in

their primary language and were subsequently debriefed by a bilingual (Spanish-English) interviewer.
Results: US Spanish subjects (N = 31) and English subjects (N = 35), stratified equally by educational level, with and without a high school degree participated in the study. The WPAI-GH item comprehension rate was 98.6% for Spanish and 99.6% for English. Response revision rates during debriefing were 1.6% for Spanish and 0.5% for English. Responses to hypothetical scenarios indicated that both language versions adequately differentiate sick time taken for health and non-health reasons and between absenteeism and presenteeism.
Conclusion: Linguistic validity of the US Spanish translation of the WPAI:GH was established among a diverse US Spanish-speaking population, including those with minimal education.
Keywords: cognitive debriefing, linguistic validation, US Spanish workers, work productivity, WPAI.

Introduction

Comprehensive evaluations of disease burden and therapeutic interventions should include measures of health-related work productivity loss, that is, absenteeism (missed work time) and presenteeism (reduced on-the-job effectiveness). Although there are several questionnaires that are useful in characterizing the impact of disease and effectiveness of treatment on work productivity [1], none of these measures has been validated in Spanish-speaking workers in the United States. In fact, there is no report in the literature concerning the validation of measures to assess the health-related work productivity loss of this population. In 2000, there were an estimated 35.3 million Hispanics in the US population (12.5%) [2], and their numbers are expected to increase to 102.6 million (24.4%) by 2050 [3]; the current Hispanic US workforce of 18.2 million is projected to increase by 10 million in 2020 [4]. Spanish is currently the dominant language for 47% of the entire Hispanic

population, with only 25% of the population being English dominant and 28% bilingual [4]. Developing a valid measure of work productivity loss for the Spanish-only population is important so that health economic studies are representative of the entire US workforce.

If assessments of work productivity for English and Spanish-speaking subjects are to be pooled or compared, it is essential that the underlying source questionnaire and the Spanish translation be conceptually equivalent (linguistic validity). Equivalence of a new translation is determined through a reiterative process of creating harmonized forward and back translations by independent translators, cognitive debriefing of subjects after questionnaire administration, and finally, psychometric testing of the translation to establish the validity, reliability, and responsiveness of the measures in the target language [5–8]. Psychometric testing of US Spanish questionnaires is frequently reported [9–11], but less attention has been paid to the crucial intermediate step of cognitive debriefing. This is an unfortunate omission because characteristics of the target population, such as lower literacy rates, lack of familiarity with questionnaires, and the need for idiomatic translations may undermine the psychometric testing [12]. When cognitive debriefing of new trans-

Address correspondence to: Mary C. Gawlicki, Corporate Translations, Inc., 77 Hartland St., East Hartford, CT 06108, USA. E-mail: mgawlicki@corptransinc.com

lations is conducted for only a handful of subjects [13,14] the resulting information may be useful for improving the translation, but the effect of pertinent demographic characteristics, such as educational level and regional variations in the target population, cannot be evaluated.

To test a measure for assessing work productivity in the US Spanish-speaking population, we selected the Work Productivity and Activity Impairment questionnaire which measures work time missed and work and activity impairment due to general health problems (WPAI:GH) or a specified health problem (WPAI:SHP) during the past 7 days [15]. The validity of the instrument has been established in a number of therapeutic areas, including allergies [16], dermatitis [17], gastroesophageal disease [18,19], nocturia [20], asthma [21], and irritable bowel syndrome (IBS) [22]. In addition, it has proven to be a useful tool in measuring the relative difference between treatment groups in clinical trials [23–26] and in subjects with and without a specific condition [27,28]. Although a US Spanish translation of the general health version of the WPAI:GH is available [29], it has not been evaluated to determine if it measures the same concepts as the original US English version and is readily understood and culturally appropriate for the population.

The objective of this study was to assess the linguistic validity of the US Spanish version of the WPAI:GH.

Materials and Methods

Subject Enrollment and Study Design

Persons who were currently employed, living in the United States, and at least 21 years of age were eligible to participate in this study. To provide a diverse study population, a convenience sample was selected in Connecticut, Florida, and New York by recruiting subjects at a laundromat, in a parking lot, mall restaurant, construction company, insurance company and through personal referrals. Subjects were recruited by a bilingual interviewer who ascertained the language the subject was most comfortable speaking. Subjects were stratified by primary language (English or Spanish) and completed all testing in the primary language. Within each language group, subjects were further stratified by education (less than 12 years of academic education vs. 12 or more years of academic education) to ensure a minimum of 15 subjects in each language group with and without the equivalent of a secondary education in the United States. Within each language group, subjects were selected to be diverse with regard to age, sex, occupation (white collar vs. blue collar), having/not having a health problem interfere with activities during the past 7 days, and for the Spanish population, country of origin. No personal information, such as names or addresses, was collected and all collected information was anonymous.

Qualified subjects completed the self-administered WPAI:GH and demographic questions. After completion of the questionnaire, a trained bilingual interviewer (A.P.) debriefed subjects on the phone, or in person, to assess the subjects' ability to paraphrase the WPAI:GH instructions and questions. For any misinterpreted item or incorrect response identified by subjects during the debriefing interview, the interviewer recorded any resulting change in the response. Two hypothetical questions were also posed to test the subjects' understanding of the need to exclude sick-day absences for nonhealth reasons from absenteeism for health reasons, and to exclude absenteeism in the calculation of productivity at work. Subjects who could not understand the task presented in the hypothetical scenarios after minimal probing were excluded from the analyses of these questions.

Questionnaire

The WPAI:GH consists of six questions that elicit the following: employment status; hours missed due to health problems; hours missed due to other reasons; hours actually worked; and two questions that measure the degree health problems affected productivity while working (presenteeism) and regular daily activities, on a scale from 0 to 10. Scores for absenteeism, presenteeism, overall work productivity loss (combined absenteeism plus presenteeism) and impairment in regular daily (nonwork) activities, such as work around the house, shopping, child care, exercising, studying, etc., are derived for the interval of the past 7 days; scores are expressed as percent of impairment/productivity loss, with higher scores indicating greater impairment. The US Spanish version was created from the US English version using standard translation procedures for questionnaires [30]. In this process, two bilingual translators created independent forward (Spanish) translations of the questionnaire, compared their results, and then created a harmonized version that best reflected the original. A third bilingual translator created a back translation (English) of the harmonized version and submitted it to the questionnaire author (M.C.R.) for review. This process was repeated until all translation issues were resolved.

After the subject paraphrased the question about missed work time due to health, the interviewer posed the following hypothetical question: "If you took a sick day, that is, called in sick, but weren't really sick, how would you have answered this question about hours missed due to health?" After paraphrasing the work productivity question, the subject was asked: "If you missed 4 days from work because you were sick, but the fifth day you went to work and had absolutely no health problem, how would you answer this question about productivity at work?"

Statistical Analysis

A descriptive analysis was performed separately for the two language groups, and by education level within each language group. An item comprehension rate was calculated for each language group by dividing the total items correctly paraphrased by the total number of items (instructions plus six questions). A response revision rate was calculated for each language group by dividing the total responses revised during debriefing by the total number of responses to the six questions.

Results

Characteristics of the Population

Table 1 shows the characteristics of the study population. A total of 66 subjects were enrolled with a primary language of English (N = 35) or Spanish (N = 31). The majority was female (54.5%) and the mean age was 40.8 years. Slightly less than half had a white-collar occupation (43.9%). The two language groups were generally comparable for these characteristics.

Average years of education was 11.7, with Spanish subjects having a lower average than English subjects (10.9 vs. 12.4) and more likely to have less than 12 years of education (48.4% vs. 42.9%). All of the Spanish subjects were born outside the United States, in one of eight Spanish-speaking countries; 88.6% of the English subjects were born in the United States.

Compared with English subjects, Spanish subjects were more likely to report a health problem interfering with activities in the past 7 days (71.0% vs. 62.9%) and reported more absenteeism (11.3% vs. 5.0%), presenteeism (26.0% vs. 20.6%), overall work productivity loss (32.9% vs. 23.5%), and daily activity impairment (34.5% vs. 24.0%), as measured by WPAI:GH scores.

Item Comprehension

The debriefing results are shown in Table 2. Of the 31 Spanish subjects, all subjects correctly paraphrased the concepts in the instructions and six questions, with the following exceptions: one subject with 4 years of education did not understand the activity limitation question until it was read out loud in the debriefing interview; one subject with 12 years of education and another with 19 years of education misinterpreted the work productivity question. Of the 34 English subjects who understood the paraphrasing task, all subjects correctly paraphrased the concepts with the following exception: one subject with 9 years of education

Table 1 Demographic, work productivity, and activity impairment characteristics of the population by primary language

Characteristic	Primary language		
	Spanish (N = 31)	English (N = 35)	Total N = 66
Age mean years (range)	39.9 (21–64)	41.5 (21–75)	40.8 (21–75)
Sex: (% females)	51.6% (16)	57.1% (20)	54.5% (36)
Occupation (% white-collar)	48.4% (15)	40.0% (14)	43.9% (29)
Education mean years (range)	10.9 (3–19)	12.4 (8–20)	11.7 (3–20)
<12 years education	48.4% (15)	42.9% (15)	45.5% (30)
3–6 years	(5)	(0)	(5)
7–8 years	(4)	(5)	(9)
9–11 years	(6)	(10)	(16)
12–16 years	(14)	(15)	(29)
17 or more years	(2)	(5)	(7)
Country of origin			
United States	0	88.6% (31)	47.0% (31)
Other, non-Spanish	0	11.4% (4)	6.1% (4)
Other, Spanish	100% (31)	0	47.0% (31)
Puerto Rico	(7)		
Dominican Republic	(6)		
Cuba	(5)		
Guatemala	(3)		
Nicaragua	(3)		
Honduras	(3)		
Peru	(2)		
Mexico	(2)		
Health problem past seven days	71.0% (22)	62.9% (22)	66.7% (44)
Absenteeism (range)	11.3% (0–100%)	5.0% (0–100%)	8.0% (0–100%)
Presenteeism (range)	26.0% (0–100%)	20.6% (0–70%)	23.2% (0–100%)
Overall work productivity loss (range)	32.9% (0–100%)	23.5% (0–100%)	27.9% (0–100%)
Activity impairment (%)	34.5% (0–100%)	24.0% (0–80%)	28.9% (0–100%)

Table 2 Number of subjects who incorrectly paraphrased concepts, revised responses or responded incorrectly to hypothetical scenarios, and item comprehension and response revision rates during debriefing of the Work Productivity and Activity Impairment Questionnaire—General Health Version (WPAI:GH) by primary language and education

	Spanish subjects Education		English subjects Education	
	<12 years (N = 15)	≥12 years (N = 16)	<12 years (N = 15)	≥12 years (N = 20)
Subjects incorrectly paraphrasing				
Instructions	0	0	0	0
Employment status	0	0	0	0
Hours missed—health	0	0	0	0
Hours missed—other reasons	0	0	0	0
Hours worked	0	0	0	0
Productivity at Work	0	2	1	0
Activity Impairment	1	0	0	0
Item comprehension rate	98.6% (214/217)		99.6% (244/245)	
Subject revising responses				
Employment status	0	0	0	1
Hours missed—health	0	0	0	0
Hours missed—other reasons	0	0	0	0
Hours worked	1	1	0	0
Productivity at Work	0	0	0	0
Activity Impairment	0	1	0	0
Response revision rate	1.6% (3/186)		0.5% (1/210)	
Subjects incorrectly responding to hypothetical scenarios	(N = 15)	(N = 16)	(N = 15)	(N = 20)
Included "sick day" used for personal reasons in hours missed due to health	1	2	1	2
	(N = 14)	(N = 15)	(N = 15)	(N = 17)
Included absenteeism in productivity at work	1	1	2	0
Included other factors in productivity at work	1	1	2	1

included absenteeism in the concept of productivity at work. Item comprehension rates were 98.6% for Spanish subjects and 99.6% for English subjects.

Response Revision

For the 31 Spanish subjects, there were no revisions to responses during the debriefing, with the following exceptions: one subject with 13 years of education revised the activity impairment response from four to six, and two subjects indicated that they had mistakenly reported hours actually worked for the prior 2-week paycheck period rather than the past 7 days. Of these two subjects, one subject with 4 years of education revised the response for hours worked from 100 to 50, resulting in no change in absenteeism, and the other with 14 years of education revised the response from 48 to 0, resulting in a change in absenteeism from 45% to 100%. For the 35 English subjects, there were no revisions to responses during the debriefing, with the following exception: one subject with 18 years of education revised her response to the employment question from not employed to employed, but answered the remaining work questions correctly. Response revision rates were 1.6% for Spanish subjects and 0.5% for English subjects.

Responses to Hypothetical Scenarios

Of the 31 Spanish subjects, one subject with 8 years of education and two subjects with 12 years of education would have included the sick day for nonhealth reasons in work time missed due to health (9.7%). Of the

35 English subjects, two subjects with 12 years of education and one subject with 11 years of education would have included the sick day (8.5%).

Among the 29 Spanish subjects who completed the hypothetical question about productivity at work, one subject with 6 years of education and another with 19 years of education would have included absenteeism in calculating the productivity response. Another two subjects considered worry about loss of income and work backlog in their calculations of productivity on the day worked. The responses for these four subjects ranged from three to eight. Among the 32 English subjects who completed the hypothetical productivity question, one subject with 9 years of education, and another with 11 years of education would have included absenteeism in calculating the response. Another three subjects considered worry about loss of income, work backlog, and employer disapproval in their calculations of productivity on the day worked. The responses for these five subjects ranged from six to nine.

Discussion

There are no measures of health-related absenteeism and presenteeism validated for use in the large and increasing Spanish-speaking population in the United States. To evaluate the usefulness of the WPAI:GH in this population, we debriefed a diverse sample of employed subjects after self-administration of the US Spanish version of the WPAI:GH and compared results

to those obtained from a comparable US English-speaking population. We found that concepts in the US Spanish WPAI:GH questionnaire were readily understandable by Spanish subjects, even among those with minimal education, and that the comprehension rate was comparable to the rate for English subjects. No errors in the US Spanish translation were found. When subjects had the opportunity to review their self-administered responses during the follow-up interview, the response revision rate was very low for both language groups, regardless of education level. Random error, rather than any education or language effect, appears to explain the reason for the few revised responses. These results are consistent with the report by Reilly et al. that the interviewer-administered WPAI provides more valid results than self-administration [15], and suggest that the low random error and high comprehension rate could be improved even further with interviewer-administration in cases where the extra effort and expense are justified.

To provide an additional test of the comparability between the US Spanish and US English versions of the WPAI:GH, we posed two hypothetical questions to subjects concerning work time missed and productivity at work. Almost all subjects in both language groups reported that they would exclude sick time for non-health reasons when answering the question about work time missed for health reasons, despite the fact that in so doing, they were admitting to their own inconsistency. This suggests that for both US Spanish and US English subjects, the WPAI:GH self-reported sick time is a more accurate and responsive measure of health-related absenteeism than data obtained from employment records where nonhealth absences are often counted as health-related. These findings are consistent with a study of the WPAI in IBS (WPAI:IBS) that found support during debriefing for 99.5% of hours missed due to IBS [21].

Responses to the hypothetical question about productivity at work indicate that some subjects in both language groups included absenteeism in their calculation of presenteeism, and others included nonhealth considerations in their calculations. Because the latter group originally paraphrased the presenteeism question correctly, we speculate that some of them may have been responding to the ambiguities inherent in the hypothetical scenario and in fact did understand the underlying concept of excluding absenteeism. This requires a modification to the scenario in future testing, for example, by inverting the scenario so that the day worked is presented first. Nevertheless, to address the inclusion of absenteeism in the response to productivity at work for some US Spanish and US English subjects, we speculate that inserting an additional instruction directly above the scale in both language versions would reinforce the intent of the question regarding productivity at work.

By including a US English-speaking comparison group and stratifying by education level, we were able to evaluate whether any observed problems with the US Spanish version of the WPAI:GH were a function of the translation itself or the result of any underlying problems in the original US English. The inclusion of hypothetical scenarios provided additional insights into the subjects' comprehension of the key concepts of absenteeism and presenteeism. We believe that this approach would be a useful methodology for establishing the linguistic validity of other health outcome questionnaires as well.

Linguistic validation of the US Spanish WPAI:GH indicates that it is a promising tool for measuring absenteeism and presenteeism in the large and increasing Spanish-speaking population in the United States. Additional investigation into its construct validity and responsiveness to clinical change is warranted.

Conclusion

Linguistic validity of the US Spanish translation of the WPAI:GH was established among a diverse US Spanish-speaking population, including those with minimal education.

References

1 Prasad M, Shih YCT, Wahlqvist P, Shikiar R. A critical review of health-related productivity measures. Value Health 2002;5:A535–6.
2 Greico EM, Cassidy RC. US Census Bureau; Overview of Race and Hispanic Origin; Census Brief. March 2001. Available from: http://www.census.gov/prod/2001pubs/c2kbr01-1.pdf [Last accessed October 6, 2005].
3 US Census Bureau News. Census Bureau Projects Tripling of Hispanic and Asian Populations in 50 Years; Non-Hispanic Whites May Drop to Half of Total Population. March 18, 2004. Available from: http://www.factfinder.census.gov/jsp/saff/SAFFInfo.jsp?_page1d=tp9_race_ethnicity. [Accessed March 2, 2006].
4 Pew Hispanic Center. Hispanics: A People on the Move. Washington DC, January 2005. Available from: http://pewhispanic.org/reports/report.php?ReportID=40 [Last accessed October 6, 2005].
5 Cull A, Sprangers M, Bjordal K, et al. EORTC Quality of Life Group Translation Procedure, (2nd ed.). Brussels: EORTC, 2002. Available from: http://www.eortc.be/home/qol/translation%20manual%202002.pdf [Last accessed October 10, 2005].
6 Hendricson WD, Russell IJ, Prihoda TJ, et al. An approach to developing a valid Spanish language translation of a health-status questionnaire. Med Care 1989;27:959–66.
7 Carlson ED. A case study in translation methodology using the Health-Promotion Lifestyle Profile II. Public Health Nurs 2000;17:61–70.

8 Mathias SD, Fifer SK, Patrick DL. Rapid translation of quality of life measures for international clinical trials: avoiding errors in the minimalist approach. Qual Life Res 1994;3:403–12.

9 Romero AA, Hardart A, Kobak W, et al. Validation of a Spanish version of the pelvic organ prolapse incontinence sexual questionnaire. Obstet Gynecol 2003;102(5 Pt 1):1000–5.

10 Froman RD, Owen SV. Validation of the Spanish Life Support Preference Questionnaire (LSPQ). J Nurs Scholarsh 2003;35:33–6.

11 Elosua R, Garcia M, Aguilar A, et al. Validation of the Minnesota leisure time physical activity questionnaire in Spanish women. Invest MARATDON Group. Med Sci Sports Exerc 2000;32:1431–7.

12 Lange JW. Methodological concerns for non-Hispanic investigators conducting research with Hispanic Americans. Res Nurs Health 2002;25:411–19.

13 Acquadro C, Lafortune L, Mear I. Quality of life in multiple sclerosis: translation in French Canadian of the MSQoL-54. Health Qual Life Outcomes 2003; 1:70.

14 Badia X, Webb SM, Prieto L, Lara N. Acromegaly quality of life questionnaire (AcroQoL). Health Qual Life Outcomes 2004;2:13.

15 Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. Pharmacoeconomics 1993;4:353–65.

16 Reilly MC, Tanner A, Meltzer EO. Work, classroom and activity impairment instruments: validation studies in allergic rhinitis. Clin Drug Invest 1996;11:278–88.

17 Reilly MC, Lavin PT, Kahler KH, Pariser DM. Validation of the Dermatology Life Quality Index (DLQI) and the Work Productivity and Activity Impairment: Chronic Hand Dermatitis Questionnaire (WPAI-ChHD) in Chronic Hand Dermatitis (ChHD). J Am Acad Dermatol 2003;48:128–30.

18 Wahlqvist P, Carlsson J, Stålhammar NO, Wiklund I. Validity of a Work Productivity and Activity Impairment Questionnaire for Patients with Symptoms of Gastro-Esophageal Reflux Disease (WPAI-GERD)—Results from a cross-sectional study. Value Health 2002;5:106–13.

19 Wahlqvist P, Guyatt G, Armstrong D, et al. Responsiveness to change and English language validation of the WPAI-GERD questionnaire—results from a Canadian study. Value Health 2004;7:A349.

20 Kobelt G, Borgstrom F, Mattiasson A. Productivity, vitality and utility in a group of healthy professionally active individuals with nocturia. Br J Urol Int 2003;9:190–5.

21 Andréasson E, Svensson K, Berggren F. The validity of the work productivity and activity impairment questionnaire for patients with asthma (WPAI-asthma): Results from a web-based study. Value Health 2003;6:780.

22 Reilly MC, Bracco A, Ricci J-F, et al. The Validity and Accuracy of the Work Productivity and Activity Impairment Questionnaire—Irritable Bowel Syndrome Version (WPAI.IBS). Aliment Pharmacol Ther 2004;20:459–67.

23 Jacobs RJ, Davidson JR, Gupta S, Meyerhoff AS. The effects of clonazepam on quality of life and work productivity in panic disorder. Am J Man Care 1997; 3:1187–96.

24 Meltzer EO, Casale TB, Nathan RA, Thompson AK. Once daily fexofenadine HCI improves quality of life and reduces work and activity impairment in patients with season allergic rhinitis. Ann Allergy Asthma Immunol 1999;83:311–17.

25 Tanner LA, Reilly M, Meltzer EO, et al. Effect of fexofenadine HCI on quality of life and work, classroom and daily activity impairment in patients with seasonal allergic rhinitis. Am J Man Care 1999; 5(Suppl.):S235–47.

26 Perrillo RP, Rothstein KD, Imperial JC, et al. Comparison of quality of life, work productivity and medical resource utilization of peginterferon alpha 2a vs. the combination of interferon alpha 2b plus ribavirin as initial treatment in patients with chronic hepatitis C. J Viral Hepatitis 2004;11:157–65.

27 Dean BB, Crawley JA, Schmitt CM, et al. The burden of illness of gastro-oesophageal reflux disease: impact on work productivity. Aliment Pharmacol Ther 2003;17:1309–17.

28 Wittchen HU, Beloch E. The impact of social phobia on quality of life. Int Clin Psycopharmacol 1996; 11(Suppl. 3):S15–23.

29 Reilly Associates. WPAI Translations. Available from: http://www.reillyassociates.net/ WPAI_Translations.html [Last accessed October 6, 2005].

30 Gawlicki MC. Procedures for Linguistic Validation of Health Status Questionnaires. Mansfield Center, CT: Corporate Translations, Inc, 2004. Available from: http://corptransinc.com/linguistic_validate/Linguistic%20Validation.pdf [Last accessed October 10, 2005].