Final Report

Developing Culturally and Linguistically Appropriate Prenatal Health Education Materials for Spanish-speaking Women
Grant number RO3 HS09836-01

October 1998 - 1999

Submitted to
Agency for Health Care Policy and Research
CareOregon
University of Washington

By
Virginia Gonzales, MSW, MPH, Ed.D
Principal Investigator
& Sandra Smith, MPH, CHES
Project Director

University of Washington
Center for Health Education and Research
Executive Summary

Developing Culturally and Linguistically Appropriate
Prenatal Health Education Materials for Spanish-speaking Women
Grant number RO3 HS09836-01
Virginia Gonzales, MSW, MPH, Ed.D & Sandra Smith, MPH, CHES
University of Washington Center for Health Education and Research

The purpose of this project is to improve the quality of prenatal care for women whose first language is not English. Our goal is to assure that women of diverse cultures and varying literacy skills have access to essential prenatal care information according to Public Health Service guidelines (PHS 1989), via a tested process for tailoring existing information.

Background
We pilot tested processes and instruments to adapt existing English language prenatal health education materials to serve the needs of Spanish-speaking women in the US. Translation, pre-testing and pilot testing directly involved learners and educators in development of materials intended for their use.

Before initiating this project, we conducted a rigorous translation and pre-testing process to discover language acceptable and persuasive to Spanish-speakers from various cultures present in the US. Through a partnership with the City of Hartford (CT) Health Department, the initial translation was pre-tested with Spanish-speakers from five Hispanic cultures present on the East Coast. A second translation team on the West Coast incorporated the Mexican perspective.

Methods
In this project, we pilot tested instruments and procedures to evaluate the translated and culturally adapted materials. Since the test materials had been pre-tested with East Coast Spanish-speakers, we conducted this pilot project with West Coast residents, primarily from Mexico. Cloze testing produced quantitative data and Reader Verification and Revision Interviews produced qualitative data. We administered these tests with 35 randomly selected, currently or recently pregnant women whose first language is Spanish and whose prenatal care was financed through CareOregon, a Medicaid managed care organization based in Portland, Oregon. Findings guided revision of the test materials and led to guidelines for adapting English health education materials to serve the needs of a typical managed care organization that is culturally and socioeconomically diverse with a range of literacy levels.

Results & Discussion
Testing showed that the tailoring process successfully produced culturally and linguistically appropriate materials from English source materials. The adaptation proved suitable for 83% of our sample of Medicaid-eligible women with average 9 years education. Half of those with six to eight years education and 80 percent of those with 9 to 12 years education found the materials easy-to-read and understand independently. We revised the materials according to our test findings. We are confident that these changes further increased the level of comprehension and the proportion for which materials are suitable. Future research should include follow-up evaluation to verify that revisions are acceptable, produce the intended remedy, and do not produce unintended effects.

Lessons Learned
Preliminary findings showed that qualitative data from personal interviews indicated higher level understanding than the quantitative data from cloze testing. This led us to a detailed investigation of the cloze test. The cloze test is a school-based test designed in 1953 to evaluate students’ literacy skills. It has been used to evaluate health education materials, and more recently to measure adult’s “health literacy” — ability to understand health related information. The test measures comprehension and other skills that are not important for health services delivery such as spelling, conjugation and vocabulary. We were concerned only with adult learners’ comprehension—how well they deciphered the meaning.
The literature indicates that cloze test results vary significantly with different test forms, that low literacy is associated with feelings of shame, and that the school-based nature of the test might induce anxiety in those who were unsuccessful in school. This is particularly significant in light of new research demonstrating that anxiety triggers release of hormones that interfere with cognitive performance.

The literature suggested modifications to remedy these problems with the cloze test, which we implemented. Most notably, we calculated a second modified score that accepted misspellings and near synonyms. While univariate and multivariate analyses showed no correlation between standard scores and any demographic factor, modified scores were significantly correlated with years of education, number of children and time in the country. Qualitative data supported the modified scoring.

We concluded that accepting synonyms produces cloze test scores that reveal ability to decipher meaning despite low literacy skills, account for previous knowledge, and may mitigate test-induced anxiety. This has implications for the practice of using the standard cloze procedure to measure "health literacy. Since the standard cloze test procedure does not account for previous knowledge or ability to decipher meaning despite low literacy skills, and does not mitigate test-induced anxiety, it may significantly underestimate adults' ability to understand health information.

Reader Verification Interviews revealed weaknesses in the materials and suggested remedies. Asking representative learners the meaning of particular terms was very useful in detecting cultural differences among Spanish speakers and subtleties in the language that could have led to misinterpretation. The interviews also revealed previously held attitudes and beliefs, and confirmed earlier research on the way that persons with limited literacy skills read and learn. Using both qualitative and quantitative data provided an effective check and balance to increase the strength of findings.

Recognizing the cultural diversity among Spanish-speakers, these findings suggest that it is possible to produce a single set of materials that is acceptable and persuasive to almost all Spanish-speakers in the US. Most importantly, this project demonstrated that it is feasible to provide equitable information to all segments of a diverse managed care population.
**ABSTRACT** (Maximum 200 words):

This project tested a protocol for adapting existing English health education materials to serve non-English-speakers. We adapted prenatal health information to serve Spanish-speaking women. Cloze testing produced quantitative data and Reader Verification Interviews produced qualitative data. We found no correlation between standard cloze test scores and any demographic variable. Modifications to the cloze test reveal ability to decipher meaning despite low literacy skills. Modified scores were significantly associated with years of education, number of children and time in the country. Qualitative data support modified scoring. Translated test materials proved easy-to-read, acceptable and persuasive for Medicaid-funded mothers averaging 9 years education.

**SUBJECT TERMS**

Unclassified

**SECURITY CLASSIFICATION OF REPORT**  
Unclassified

**SECURITY CLASSIFICATION OF THIS PAGE**  
Unclassified

**SECURITY CLASSIFICATION OF ABSTRACT**  
Unclassified

**LIMITATION OF ABSTRACT**  
Unclassified

**REPORT DOCUMENTATION PAGE**

**NSN**  
7540-01-280-5500

**Releasable**

To public; available from NTIS.

**REPORT TYPE AND DATES COVERED**

abstract, executive summary and final report, 09/30/98-09/29/99

**FUNDING NUMBERS**

Grant HS 09836

**PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**

University of Washington  
Seattle, Washington

**SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**

DHHS, PHS  
Agency for Health Care Policy and Research (AHCPR)  
Center for Research Dissemination and Liaison  
Executive Office Building, Suite 501, 2101 East Jefferson Street  
Rockville, Maryland 20852

**SUPPORTING / MONITORING AGENCY REPORT NUMBER**

AHCPR No. 00-11

**DISTRIBUTION / AVAILABILITY STATEMENT**

Releasable to the public; available from NTIS.

**ABSTRACT**

This project tested a protocol for adapting existing English health education materials to serve non-English-speakers. We adapted prenatal health information to serve Spanish-speaking women. Cloze testing produced quantitative data and Reader Verification Interviews produced qualitative data. We found no correlation between standard cloze test scores and any demographic variable. Modifications to the cloze test reveal ability to decipher meaning despite low literacy skills. Modified scores were significantly associated with years of education, number of children and time in the country. Qualitative data support modified scoring. Translated test materials proved easy-to-read, acceptable and persuasive for Medicaid-funded mothers averaging 9 years education.

**SUBJECT TERMS**

Unclassified

**SECURITY CLASSIFICATION OF REPORT**  
Unclassified

**SECURITY CLASSIFICATION OF THIS PAGE**  
Unclassified

**SECURITY CLASSIFICATION OF ABSTRACT**  
Unclassified

**LIMITATION OF ABSTRACT**  
Unclassified

ISN 7540-01-280-5500

PROTECTED UNDER INTERNATIONAL COPYRIGHT  
ALL RIGHTS RESERVED.  
NATIONAL TECHNICAL INFORMATION SERVICE  
U.S. DEPARTMENT OF COMMERCE
Abstract

This project tested a protocol for adapting existing English health education materials to serve non-English-speakers. We adapted prenatal health information to serve Spanish-speaking women. Cloze testing produced quantitative data and Reader Verification Interviews produced qualitative data. We found no correlation between standard cloze test scores and any demographic variable. Modifications to the cloze test reveal ability to decipher meaning despite low literacy skills. Modified scores were significantly associated with years of education, number of children and time in the country. Qualitative data support modified scoring. Translated test materials proved easy-to-read, acceptable and persuasive for Medicaid-funded mothers averaging 9 years education.
Final Report

Developing Culturally and Linguistically Appropriate Prenatal Health Education Materials for Spanish-speaking Women
Grant number RO3 HS09836-01

Statement of Purpose
The purpose of this project is to improve the quality of prenatal care for women whose first language is not English. Our goal is to assure that women of diverse cultures and varying literacy skills have access to essential prenatal care information according to Public Health Service guidelines (PHS 1989), via a tested process for tailoring information. Specifically, we aim to increase access to key messages linked to positive birth outcomes (Kogan1994, Davis 1987, Libbus 1991). This first step pilots a process for adapting existing English materials to serve the perinatal information needs of Spanish speaking women in the United States. The results of this project are culturally and linguistically appropriate learning materials for Spanish-speakers and guidelines for adapting English health education material to serve the needs of diverse managed care populations.

Methods

Study Population
From January to August 1999, we recruited and interviewed project participants in Portland, Oregon. We randomly selected potential participants from the Spanish-speaking segment of women who obtained prenatal care through CareOregon, a Portland-area Medicaid managed care organization. CareOregon generated from its database a list of women who met the eligibility criteria. Using the S-Plus statistical package, we assigned women a random number and contacted them in random order.

Eligible women were aged 18-45 years and spoke Spanish as their first language. Each obtained at least one prenatal care visit through CareOregon during 1997 to 1999, attained at least a sixth grade education, and resided in the Portland area.

Participants ranged in age from 18 to 38, with an average age of 27 years. Most of the women (83%) were married and not employed (74%) at the time of the interview. Most were born in Mexico (83%); 12% were born in the United States.

Nearly three quarters of the women (71%) identified themselves as Latino or Hispanic and 10% identified themselves as "descendiente de mejicano (descendents of Mexicans). Over half the women (57%) have lived in the country ten or more years. On average, the women have been US residents for 7 years, up to 23 years. Over half (51%) completed six to nine years of schooling and 46% completed 10-12 years. Average educational achievement was 9 years. Table 1 (Appendix) shows demographic characteristics of the study sample.

Test Materials
We selected source materials already rigorously reviewed for scientific accuracy and purposely developed for a typical managed care population that is socioeconomically and culturally diverse with a wide range of literacy skills. The materials, Beginnings: A Practical Guide through Your Pregnancy (third edition), have been accepted nationally by the medical community and managed care organizations since 1989. Readability ratings for the materials are 4th grade on the Fry and Flesch-Kincaid scales and 88.1 (easy) on the Flesch Reading Ease Index. In previous testing, both college-educated women and those with less than ninth grade education reported extreme satisfaction with these materials (Smith 1998).

Translation
While translation and cultural tailoring of the test materials was not specifically part of this project, the process and quality of translation is a significant factor in suitability and effectiveness of materials. Straight translation ignores differences among cultural groups and subtleties in the language.
To achieve our goal of producing materials suitable for almost all Spanish-speakers in the US, we formed a partnership with the staff of the perinatal outreach program of the City of Hartford (CT) Health Department. This team of medical professionals and community health workers are Spanish speakers from Cuba, Puerto Rico, Guatemala, Columbia, and Peru serving clients from the same countries. This team translated the source materials using terms familiar in each of their Spanish dialects and then incorporated revisions suggested by their respective clients. A second translation team in Seattle, WA incorporated the Mexican perspective and edited the text as necessary to fit the pages. The team leaders resolved any remaining differences. Selections from the resulting text made up the test materials. Since the text had been pre-tested with women from the Spanish-speaking cultures present in the eastern US, our testers for this project were West Coast residents, primarily from Mexico.

Recruitment

We attempted to contact 234 women in random sequence. Of these, 92 could not be reached due to wrong, missing, or disconnected telephone numbers, and 40 either did not answer or did not respond to messages after three attempts and a postcard. Of the 102 women successfully contacted, 27 were not eligible, 23 were not interested, and nine had scheduling problems. We scheduled 43 interviews; eight did not show. The principle investigator interviewed 35 women, a suitable number for this type of formative research (Doak and Doak 1996).

Screening and Interviews

A Spanish-English bilingual employee of CareOregon called potential participants in the assigned random order. If a woman was ineligible, declined to participate, or could not be reached, the caller contacted the next person on the list. She selected names from the master list as needed until at least 35 participants were recruited and interviewed.

Fifteen-minute telephone screenings included an explanation of the participant selection process and the purpose of the study. The caller confirmed eligibility and then explained compensation and incentives. For each eligible and interested woman, the caller scheduled a one-hour face-to-face interview to be conducted at the most convenient of four clinic locations or in her home. We sent a confirmation letter to each woman who scheduled an appointment.

The project paid a $25 incentive to each participant who completed the interview. In addition, the project offered participants bus tokens to and from the interview site, on-site childcare, light refreshments and the materials that she reviewed.

Data Collection Procedures and Instruments

To evaluate the translated test materials, we used two instruments: the cloze procedure (Taylor 1953) and Learner Verification and Revision Interviews (Doak and Doak 1996). The principal investigator and a CareOregon staff person conducted the one-hour interviews in Spanish. Each interview consisted of an introduction (5 minutes); cloze test (20-30 minutes), break and refreshments (5 minutes), learner verification and revision interview (20-30 minutes).

Introduction to the interview

The interviewers explained the purpose of the study and confidentiality protections and obtained participants’ permission to audio record interviews. To protect participants’ confidentiality, identification numbers replaced names on all study documents and audiotapes. Only research staff had access to study documents and the data. Spanish-speaking project staff transcribed the taped interviews in Spanish and the investigator translated them to English for analysis.

The cloze test

The cloze test is a self-administered paper and pencil test designed to evaluate students’ comprehension of reading material. It is validated in college students and the military (Foltz 1998), and is commonly used to acquire and test foreign language proficiency. Validity has not been established for older
adults (Foltz 1998), or Spanish-speakers. Doak and Doak (1996) suggest the cloze test is useful in evaluating the suitability of health education materials for readers with a minimum 6th grade education. Recently, Weiss (1992) and others used the procedure to test patients’ “health literacy”.

The cloze test is not timed, but usually takes 20 to 30 minutes. To create the instrument, every fifth word is deleted from a sample of text. The task is to read the sample text, and then, without referring to the text, fill in the blanks with exact replacements. The cloze test measures comprehension by testing how much knowledge the reader obtained from information surrounding the blanks. More importantly for this project, the procedure reveals terms and concepts that readers find difficult, unfamiliar, or unacceptable, and “wrong” responses suggest more appropriate language.

There are difficulties with the cloze procedure in adults in addition to its school-based design. Research shows that different forms of the test produce significantly different results and that at least five different forms should be used in each evaluation to assure accuracy (Elley 1979). Researchers acknowledge this rarely is feasible. To increase validity of our cloze testing with a single form, we followed Elley’s suggestion to ensure that the proportion of nouns tested equals approximately 20% of the total number of nouns in the test passage. Busselman and Holcomb (1994) conducted cloze testing with a study population similar to ours. We followed their suggestion to reduce participants’ frustration by using word-length blanks instead of uniform-length blanks. Rush and Klare (1978) validated this modification.

To reduce participants’ anxiety, the interviewer emphasized that we were testing the materials, not the participant, as suggested by Doak and Doak (1996). Reducing anxiety is significant since Newcomer (1999) demonstrated that anxiety triggers secretion of hormones that reduce cognitive ability. The school-based cloze test is likely to induce anxiety in those who were unsuccessful in school and those embarrassed by their low literacy (Parikh 1996). In addition, testing by representatives of government or institutions may induce anxiety in those for whom immigration status is a factor.

A number of cloze test scoring procedures are documented in the literature. Bormuth established 40% (20 out of 50) correct as the threshold for concluding that material is within the student’s comprehension (Elley 1979). Raggett (1979) suggested 44% (22 of 50) as the threshold for comprehension and a score of ≥ 57 to indicate full independent understanding. According to Busselman and Holcomb (1994) scores of 44% to 56% equate to 75% comprehension while scores ≥ 57% equate to 90% text comprehension. Doak and Doak (1996) suggest that scores under 40% show the reader finds the text frustrating, while scores of 40 to 60% show text will be useful to readers although they may need assistance to comprehend fully, and scores above 60 show easy reading and full understanding. For this project, we adopted the following scoring:

<20 correct (<40%): materials are too difficult for the reader.

≥20 correct (≥40%): materials are useful. Assistance may be required for complete understanding.

≥29 (≥57%): materials are easy to understand, suitable for independent learning.

Standard scoring accepts only exact replacements for the words deleted on the test instrument. For the purposes of this project, we were not concerned with some literacy skills that this standard scoring is designed to measure, such as spelling, conjugation and vocabulary, but only with understanding. Therefore, we calculated a second modified score that disregarded misspellings and accepted near synonyms. Both of these variations are documented in the literature (Holcomb 1978, Elley 1979). In a review of the literature, Jongsma (1970) documented higher reliability with synonyms accepted. Table 2 (Appendix) shows modifications to the cloze procedure with reasons and evidence for each modification.

For each blank on the cloze test, the research team entered two variables into a database: 1) the participants’ written response, and 2) a Yes/No indicating whether the response was an exact replacement. From this data, we produced a summary of responses and calculated a standard score for each woman. A translation team leader reviewed the summary of responses and identified apparent misspellings and acceptable synonyms. We then entered a third variable into the database indicating whether each response was acceptable as a synonym and from this calculated the modified score. Using SPSS for Windows Version 9.0, we analyzed the test scores by various demographic variables (e.g., years of education, reading frequency, employment, parity, number of years in the US, age) to see if and how these factors affected comprehension.
Learner Verification and Revision Interview

For this second part of the interview, we gave each participant a prototype of the first booklet in a six-part series. The women viewed illustrations and read sections of the test booklet that we suspected might be misinterpreted, irrelevant, or unacceptable. Then the principal investigator conducted the Learner Verification and Revision Interview using a qualitative open-ended questionnaire designed to uncover specific content and format features that the participant did not understand or accept, and to produce solutions. The interviewer audio taped each session. Working with transcribed and translated manuscripts, the research team categorized responses to each question according to content and produced a summary.

Results & Discussion

Cloze Testing

For the purposes of this project, we were not concerned with some literacy skills that the cloze test is designed to measure, such as spelling, conjugation and vocabulary. Rather we were concerned only with understanding. Therefore, we calculated a second modified score that disregarded misspellings and accepted near synonyms. Both of these variations are documented in the literature (Holcomb 1978, Elley 1979). Although Jongsma (1970) concluded that the standard scoring method is more efficient, he documented higher reliability with synonyms counted.

Other researchers found that accepting or not accepting synonyms made little difference since modified scores correlated strongly with standard scoring results. Elley (1979) reported that accepting synonyms increased high school students' number correct by 3 or 4 in scientific material and by only 1 or 2 in prose. In our sample of Medicaid mothers long out of school with average 9 years of education, accepting synonyms resulted in a typical gain of 6 or 7 with gains up to 11. These gains demonstrated that most participants were able to decipher the meaning although their skills were insufficient to reproduce the words precisely. Thus, the modified scoring differentiated those who gained knowledge from the text despite low literacy skills from those who could not make use of the information.

For example, this phrase from the cloze test is translated from If someone is hurting you, get help from your doctor or midwife: Si alguien la esta ________, obtenga ayuda de su ______ o obstetra. The correct replacement for the first blank is lastimando. Only 20% of testers gave this response. However, an additional 66% responded with synonyms equivalent to hitting-, mistreating-, harassing-, molesting-, abusing-, and acting violently toward you. These synonymous responses showed that 86% of the women understood the meaning and cued us to change the wording to their most common response. Similarly, we accepted terms such as médico as synonyms for doctor.

Overall, the average standard cloze test score was 24 correct or 48% compared to the modified average of 30 correct or 60%. Standard scoring suggested the materials might be too difficult for 29% of the women. With synonyms accepted and spelling disregarded, the scores show the materials are understandable and useful for all but 18%.

Accepting synonyms, the average score for those with 6 to 8 years education increased from 37%—slightly below the threshold for comprehension, to 49%—well above the threshold. The average for those with 9 to 10 years schooling increased from 53% to 64%; and those with 11-12 years education increased from 55% to 69%. These modified averages are well above the 57% indicator of full independent comprehension. Table 3 below shows the mean standard and modified cloze test scores by education level.

<table>
<thead>
<tr>
<th>Years Education</th>
<th>6 - 8 (N=14)</th>
<th>9 - 10 (N=5)</th>
<th>11 - 12 (N=15)</th>
<th>Overall (N=34)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score &lt;40%</td>
<td>50%</td>
<td>20%</td>
<td>13%</td>
<td>29%</td>
</tr>
<tr>
<td>40-56%</td>
<td>43%</td>
<td>20%</td>
<td>7%</td>
<td>35%</td>
</tr>
<tr>
<td>≥57%</td>
<td>7%</td>
<td>60%</td>
<td>80%</td>
<td>68%</td>
</tr>
</tbody>
</table>

Standard scores accept only exact replacements of words deleted from the text. Modified scores disregarded spelling errors and accepted synonyms.
These findings indicate that the test materials are suitable for 83% of women in this population of Medicaid eligible Spanish-speaking women. Half of those with 6 to 8 years education and 80% of those with 9 or more years found the materials easy to read and understand. About 18% do not have sufficient literacy skills to use materials written at a 3rd grade level. These persons will need face-to-face teaching aided by image-only learning materials that convey the key messages without words.

Image-only materials are urgently needed by thousands of immigrants and refugees who may be proficient in a language for which no materials exist, and by those who, like 18% of our participants, have very limited literacy skills. Development of image-only materials is a logical next step toward assuring that women of diverse cultures and varying literacy skills have access to essential prenatal care information according to Public Health Service guidelines (PHS 1989).

Demographic Factors & Cloze Test Scores

Years of Education

We performed univariate and multivariate analysis to determine effects of demographic factors on standard and modified cloze test scores. A cluster of recent articles report use of cloze testing to measure patients’ “health literacy” (Rutledge & Donaldson 1998), which Baker (1998) described as ability to understand health related information. In these reports, educational achievement is positively associated with test scores. However, the strength of the association varies widely across samples. Since the cloze test is a school-based measure, we anticipated that scores should reflect educational achievement. In this study, we found no significant correlation between standard cloze scores and years of education (adjusted R Squared = .099, p = .360). However, modified scores were significantly associated with years of education (adjusted R Squared = .802, p = .000).

Even this strong correlation between educational achievement and reading comprehension should be interpreted with caution. Years of education did not reliably predict individual scores. Notably, both the highest score — 35 correct (42 modified) — and the lowest score — 6 correct (8 modified) — were achieved by women who obtained 9 years education in Mexico. For those with 6 to 8 years education, the range of scores was 6 correct (7 modified) to 32 (41 modified). For those with 11-12 years education, the range was 13 correct (18 modified) to 34 (39 modified). These data are consistent with previous findings that most Americans read 3 to 5 grades below their educational achievement, and education is not a reliable predictor of individual literacy skills (Doak 1996, Rutledge 1998). Table 4 (Appendix) shows standard and modified cloze scores by years of education.

Reading Frequency

We anticipated that cloze scores would reflect reading frequency since, like any skill, literacy skills improve with practice. Overland (1993) found that reading frequency consistently predicted cloze scores of older English-speaking diabetic patients with average 10–11th grade education. We found no significant correlation between reading frequency and standard cloze scores (adjusted R squared = .324, p = .889) or modified scores (adjusted R squared = .314, p = .165). In this study, the woman who achieved the highest cloze score reported that she “never” reads, while a woman who scored below the threshold of comprehension said she reads every day. This suggests a need for further research.

Employment and Age

Since low literacy severely limits employment opportunities, we expected cloze scores to be associated with employment. Only those women who scored in the top half of the sample were employed. However, employment did not explain differences in either standard or modified scores. Age was not significantly associated with either standard or modified scores.

Years in the US

We reasoned that familiarity with the healthcare system and prenatal care would increase with time in the country, and comprehension scores would reflect this familiarity. Years in the US had no significant
effect on standard cloze test scores (adjusted R squared = .186, p = .246). Modified scores show a significant correlation with years in the country (adjusted R squared = .513, p = .041). This finding further suggests that accepting synonyms may more accurately reflect previous knowledge and ability to decipher meaning despite low literacy skills.

Notably, although the study participants have resided in the US on average 7 years and up to 23 years, all but one remain monolingual. The US Department of Health and Human Services (1995) estimates that more than 32 million US residents speak a language other than English at home. Persons not proficient in English face substantial communication barriers at almost every level of the health care system (DHHS 1993). These findings along with increasing immigration suggest an urgent need for easy-to-read health information not only for refugees and recent immigrants but also for an established and growing population of Spanish-speakers and other non-English speakers.

Number of Children

We reasoned that "experienced" mothers would have more previous knowledge of pregnancy and prenatal care so that cloze scores would increase with parity. Correlation between standard scores and number of children was non-significant (adjusted R squared = .347, p = .092). However, with modified scoring, the number of children explained over 50% of the difference in cloze scores at a significance level of .04 (adjusted R squared = .515, p = .041). This suggests that accepting synonyms may produce scores that more accurately reflect previous knowledge and ability to decipher meaning despite low literacy skills.

Implications for Practice

These findings have implications for the practice of using the cloze test to measure individuals’ “health literacy.” Since standard scoring does not reveal ability to decipher meaning despite low literacy skills and does not reflect previous knowledge or mitigate test-induced anxiety, it may significantly underestimate patients’ ability to understand health related information, particularly information that is directly applicable and important to them.

Reader Verification and Revision Interviews

The Reader Verification and Revision Interview process uncovers specific content or format features that learners do not understand or do not accept, and produces remedies (Doak and Doak 1996). We evaluated five elements to determine the likely influence of the test materials: attraction, comprehension, acceptability, self-efficacy and persuasion.

Attraction

The first communication need is to attract the intended learner and carry her into the message. Without strong attraction, there is no chance to influence the learner. By looking at the cover, learners need to understand the purpose of the material and see how it applies to them. A dilemma for materials designers is that a colorful highly designed page is attractive to readers, but interferes with comprehension (Wheldon 1995). Accordingly, we used a colorful cover with simply designed black and white interior pages. About three-quarters (71%) of participants accurately described the cover art as depicting a mother with child. Several specifically mentioned that the colors attracted them. Other testers interpreted the art more generally and literally as a Hispanic woman or simply a person. All but two testers (94%) said they would pick up the test material and read it. One said she would not because the words are too small, although she later said she would like to receive additional booklets in the series. The second felt no need for the information since she recently gave birth. All testers accurately described the subject matter. From these findings we concluded that the cover art is attractive.

Acceptability

To avoid unfamiliar words and to discover terms that would convey the same meaning to speakers of various Spanish dialects, we asked participants what selected terms mean to them. This was particularly useful in determining what to call prenatal care providers who are not physicians. Only 8% of testers
correctly defined *midwife*, and none accurately understood the term *obstetriz*, which is common in Peru. We chose the term *partera* since it is familiar and well understood by participants on both coasts.

While 31% correctly defined the term *self-care*, responses throughout the interview indicated that participants understood and embraced the concept, so we elected to introduce the term on the first page and use it throughout the materials. We retained the literal translation of *birth defects* since all the women defined it accurately. Since a few associated defects with mothers' actions, the text emphasizes that "things do go wrong even when everyone does everything right."

We asked directly how we should talk about urination since the correct term is rarely used in conversational English. Perhaps reflecting the more formal nature of Spanish, two-thirds of the women preferred the correct term (*urinar*), while one-third suggested the equivalent of *go to the bathroom*, which appears in the English source materials. We changed the text to reflect the majority opinion.

In the process of pretesting the translation with clients of the Hartford (CT) Health Department, the East Coast translation team noted that many clients preferred the less formal address *tu* over the more formal and proper *usted*. This led us to include a question about preference in our West Coast interviews. The women were equally divided in their preference. One-third (34%) preferred *usted* and one-third (31%) preferred *tu*. Since the remaining third of the testers said either form is fine, and since purchasers said they would not distribute materials using *tu*, we stayed with the more formal and respectful *usted*.

A major theme in the source materials is *keys to a healthy baby* – the key messages which research links to birth outcomes. The theme is carried throughout the source materials and a key symbol draws attention to key messages wherever they appear. This common metaphor does not come across in translation. About one-third of the women noted that the key symbol indicates important points, but only 6% understood the metaphor. The East Coast translation team agreed that the key was only somewhat meaningful to their clients. They suggested a happy face or light bulb would be stronger symbols.

**Comprehension, self-efficacy and persuasion**

Comprehension is critical for persons with low literacy since they can access fewer information resources than highly literate learners. We asked women to read the keys to healthy baby and then name healthy things to do during pregnancy. All successfully named several key healthful behaviors, and when asked to name 'things not to do' 86% correctly named on average, 2 or 3 of the key messages. Smoking was most frequently mentioned (75%) with drugs and alcohol following closely. It was instructive that 29% mentioned avoiding upset or maintaining tranquility. Although this concept appears throughout the materials, it is not included in the key messages. Similarly, 20% said 'don't lift heavy things', a topic that is not addressed in the test materials. These statements reveal previously held beliefs.

To judge self-efficacy—women's feelings of confidence that they can do what is required to have a healthy baby (Bandura 1986)—we asked which key message seemed most difficult for them to act on. The single most frequent response (37%) was *None*, an indication of high self-efficacy. Of those who found healthy behaviors challenging, 68% said the most difficult advice to follow is *Eat well or Gain weight*, a reflection of poverty, and perhaps of an American social preference for a slim figure. All said they would make their most difficult changes for the baby and many seemed surprised that we would ask. This finding is in accord with previous studies showing that mothers who recall of advice to gain weight during pregnancy have a significantly reduced chance of delivering a low birth weight infant (Kogan 1994).

After reading a list titled *Warning Signs*, 100% of the women named on average 3 or 4 reasons to call to the doctor, demonstrating good comprehension. Four women each added one item of previous knowledge: rashes, "excessive exercise", no fetal movement, and diarrhea. When asked what they would do if they experienced one of the warning signs, 80% said they would call the doctor or the emergency number (911) as stated in the instruction. Two women said they would call the hospital or clinic and three would go to the hospital. If the women suspected a problem but were not sure that it warranted a call to the doctor, 80% said correctly that they would call anyway, and three would go to the doctor. Three others said they would think about it. This shows that the test materials are likely to increase women's awareness of warning signs and influence them to trigger intervention in a timely manner.

All the women correctly described the food log and its purpose, and all said they would use it. Each successfully named nutritious foods from the list and described a healthy lunch that they would prepare for
themselves. All but one tester correctly described a list of ‘empty foods’ as foods to avoid and the majority recognized one or more items on the list that they eat regularly. Only one woman said she would not give up those items for the baby’s health, demonstrating that the information is persuasive and most of the women have strong self-efficacy regarding nutrition.

After reading an instruction on preventing and self-treating headaches, all the women stated one or more suggestions from the instruction and named a step they would take personally. They understood the instruction to avoid medications except Tylenol, with one quarter specifically stating that Tylenol is safe. We changed the instruction that originally read *Take a warm—not hot—bath* to simply *Take a warm bath*, because 11% specifically stated they would take a hot bath. This confirms Doaks’ (1996) observation that low literacy readers typically read one word at a time and can forget the beginning of a sentence by the time they work through to the end.

We showed the testers an illustration of a pregnant woman with Asian features demonstrating proper seatbelt use. While 40% said the pictured woman did not look like anyone they know, several mentioned that she is pregnant like them, and all agreed that the illustration was “OK to show how to wear a seatbelt”. All said they would wear the seatbelt as shown, demonstrating that the message is persuasive.

While the majority, (86%) correctly described two series of illustrations depicting the development of fetal hands and feet, a few thought the pictures showed defects until prompted by the interviewer to read the caption. While 83% accurately described a photo of a couple in discussion with a midwife or doctor, a few thought the woman’s hand gesture showed surprise or anger. Two women did not see the caption, so we made it more prominent.

No participant was offended by anything in the test materials, and all said they would like to receive the others booklets in the series. Participants’ comments shown in Table 5 (Appendix) demonstrate how easy-to-read health education materials fill gaps in women’s knowledge and enable them to make healthful changes that they are eager to make to assure their children’s health.

**Summary and Conclusions**

We pilot tested processes and instruments to adapt existing English language prenatal education materials to serve the needs of Spanish-speaking women in the US. Translation, pre-testing and pilot testing directly involved learners and educators in development of materials intended for their use. Cloze testing produced quantitative data and Reader Verification and Revision Interviews produced qualitative data.

We modified the cloze test procedure to reduce frustration and anxiety and calculated a modified score that accepted misspellings and synonyms. We analyzed effects of demographic factors on both standard and modified cloze test scores. We found no correlation between standard scores and any demographic factor. Modified scores were significantly correlated with years of education, number of children and time in the country. Qualitative data indicate the women understood the information at a higher level than the standard scores suggest and support modified scoring.

We conclude that these modifications to the cloze procedure take into account previous knowledge, reveal ability to decipher the meaning despite low literacy skills, and may mitigate test-induced anxiety. The standard cloze procedure is designed for students and may underestimate adults’ entry-level knowledge and their ability to comprehend health information that is applicable and important to them.

The test materials proved suitable for 83% of women in this population of Medicaid eligible Spanish-speaking women. Half of those with 6 to 8 years education and 80% of those with 9 or more years found the materials easy to read and independently understand. About 18% did not have sufficient literacy skills to use materials written at a 3rd grade level. We recommend development of image-only materials to meet an urgent need for teaching/learning aids for those with very limited literacy skills and those who may be proficient in languages in which materials are not available.

In this pilot project, the process we tested successfully adapted English language materials to serve the needs of most Spanish-speakers in the US. In addition, the process revealed weaknesses in the materials and produced remedies which we applied to the test booklet and which guided translation and adaptation of the series. We are confident that these improvements further increased comprehension levels and the proportion of learners for which materials are useful. Future research should include retesting to judge the effects of improvements.
Recognizing the diversity among Spanish-speakers, this project demonstrated that it is feasible to produce a single set of materials that is acceptable and persuasive to almost all Spanish-speakers in the US. More importantly, this project showed that it is feasible to provide equitable health information to all segments of a diverse managed care population.

The following guidelines are drawn from lessons we learned in the course of this project. They will assist efforts to adapt existing English materials to serve the needs of non-English speaking segments of diverse managed care populations.

**Guidelines for Adapting Existing Materials to Serve a Diverse Population**

**Materials**
- Select source materials that are demonstrated suitable and effective for their intended audience. Consider reading level and content and design factors that affect comprehension. SAM, the Suitability Assessment of Materials (Doak and Doak 1996) is a validated and scored instrument useful in evaluating source materials.
- Recognize differences among cultural groups and subtleties in the language. Avoid straight translation.
- Involve intended learners and educators in the process of translation and testing. Ensure that testers represent the mix of cultures and dialects present in the intended audience.

**Project Management**
- Increase the budget to allow for the extra time, resources and expertise that a bilingual, cross-cultural project requires.
- Select and thoroughly train professional bilingual staff with ability to establish rapport with members of the intended audience. If resource limitations require reliance on volunteers, emphasize training and communication and allow extra time to complete the project.
- Use multiple efforts to recruit participants. Face-to-face recruitment efforts will be most successful. Ideally, interviewers will be onsite and enjoy collaboration with clinic staff. The presence of two interviewers eliminates the need to make participants wait.
- When participants live in rural areas or have children, bus tokens cannot solve transportation issues. Interviews combined with scheduled clinic visits and in-home interviews are important options.
- Participants should be compensated for their time and cultural expertise.
- Include in the planning and budget follow-up evaluation to verify that revisions are acceptable, produce the intended remedy, and do not produce unintended effects.

**Instruments**
- The cloze test reveals words that are unfamiliar or subject to misinterpretation and suggests more appropriate terms. The procedure may be difficult for persons with limited literacy skills and is not useful for persons with less than six years education. The cloze test should be used to evaluate materials and not individuals. Scoring should be modified to allow misspellings and synonyms. This modification reveals testers' ability to decipher meaning despite low literacy skills and differentiates those who cannot make use of the material.
- Include demographic questions to evaluate the effect of previous knowledge. For example, in this study a question about the birthplace of the women's children might have produced additional findings about the effect of previous knowledge of prenatal care on comprehension.
- The Reader Verification and Revision Interview process reveals previously held beliefs and previous knowledge as well as concepts that do not come across in translation. The interview process uncovers and suggests remedies to elements that may be misinterpreted. Asking the meaning of specific terms is useful. Allow time to pretest and revise the questionnaire to assure that all questions produce actionable data.
Collecting both qualitative and quantitative data provides a check and balance that helps assure accurate interpretation of findings. Of the two, the qualitative data is more reliable and a richer source of actionable information about the suitability of foreign-language materials adapted from existing English materials.

Bottom Line: Learners have the last word.
References


Busselman KM and Holcomb CA (1994) Reading skill and comprehension of the Dietary Guidelines by WIC participants. *Journal of the American Dietetic Association* (94) 622-625

Davis M. Akridge K. The effects of promoting intrauterine attachment in primaparas on post delivery attachment. *JOGNN* 16(6) 1987 pp 430-7


Rush RJ, Klare GR. (1978) Reopening the cloze blank issue. *Journal of Reading Behavior* (10) 208-10


US Department of Health and Human Services, Office for Civil Rights (1995) *Limited English Proficiency as a Barrier to Health & Social Services*, Macro International:


Wheildon C. (1995)*Type and Layout: How typography and design can get your message across or get in the way*. Berkeley, CA, Strathmoor Press
Table 1 Demographic Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Number (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age Range 18-38 Mean 27.3</td>
<td>35 (100)</td>
</tr>
<tr>
<td>Years in Oregon</td>
<td></td>
</tr>
<tr>
<td>&lt; 1</td>
<td>2 (.6)</td>
</tr>
<tr>
<td>1-6</td>
<td>12 (34)</td>
</tr>
<tr>
<td>7-13</td>
<td>21 (60)</td>
</tr>
<tr>
<td>Years in US</td>
<td></td>
</tr>
<tr>
<td>&lt; 1</td>
<td>1 (3)</td>
</tr>
<tr>
<td>3-10</td>
<td>22 (63)</td>
</tr>
<tr>
<td>11-15</td>
<td>9 (26)</td>
</tr>
<tr>
<td>20-23</td>
<td>3 (9)</td>
</tr>
<tr>
<td>Country of origin</td>
<td></td>
</tr>
<tr>
<td>Guatemala</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Mexico</td>
<td>29 (83)</td>
</tr>
<tr>
<td>Panama</td>
<td>1 (3)</td>
</tr>
<tr>
<td>US</td>
<td>4 (12)</td>
</tr>
<tr>
<td>Ethnic identity</td>
<td></td>
</tr>
<tr>
<td>Latino/Hispanic</td>
<td>25 (71)</td>
</tr>
<tr>
<td>Mejicano</td>
<td>10 (29)</td>
</tr>
<tr>
<td>Where educated</td>
<td></td>
</tr>
<tr>
<td>Guatemala</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Mexico</td>
<td>28 (80)</td>
</tr>
<tr>
<td>US</td>
<td>6 (17)</td>
</tr>
<tr>
<td>Years of education</td>
<td></td>
</tr>
<tr>
<td>6-8</td>
<td>14 (40)</td>
</tr>
<tr>
<td>9-10</td>
<td>5 (14)</td>
</tr>
<tr>
<td>11-12</td>
<td>15 (43)</td>
</tr>
<tr>
<td>Missing</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Reading frequency</td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Once in awhile</td>
<td>1 (3)</td>
</tr>
<tr>
<td>1 month</td>
<td>7 (20)</td>
</tr>
<tr>
<td>2-3 month</td>
<td>4 (11)</td>
</tr>
<tr>
<td>4 x month</td>
<td>13 (37)</td>
</tr>
<tr>
<td>8-12 x month</td>
<td>6 (17)</td>
</tr>
<tr>
<td>30 x month</td>
<td>3 (9)</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>29 (83)</td>
</tr>
<tr>
<td>Divorced/Separated</td>
<td>3 (9)</td>
</tr>
<tr>
<td>Single</td>
<td>3 (9)</td>
</tr>
<tr>
<td>Currently pregnant</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>8 (23)</td>
</tr>
<tr>
<td>No</td>
<td>27 (77)</td>
</tr>
<tr>
<td>Number of children</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1 (3)</td>
</tr>
<tr>
<td>1-2</td>
<td>19 (54)</td>
</tr>
<tr>
<td>3-4</td>
<td>16 (46)</td>
</tr>
<tr>
<td>Employed</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>26 (74)</td>
</tr>
<tr>
<td>Yes</td>
<td>9 (26)</td>
</tr>
</tbody>
</table>

Note: Figures may exceed 100% due to rounding
Table 2 Modifications to the Cloze Procedure

<table>
<thead>
<tr>
<th>Modification</th>
<th>Purpose</th>
<th>Evidence/Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nouns deleted equal 20% of nouns in the test passage</td>
<td>Eliminate major cause of variability between forms</td>
<td>Elley 1979</td>
</tr>
<tr>
<td>Word-length blanks</td>
<td>Reduce frustration</td>
<td>Busselman &amp; Holcomb 1994</td>
</tr>
<tr>
<td>“We are testing the materials, not you”</td>
<td>Reduce anxiety</td>
<td>Rush &amp; Kates 1978</td>
</tr>
<tr>
<td>Disregard misspellings</td>
<td>Reveal ability to decipher meaning despite low literacy</td>
<td>Holcomb 1978</td>
</tr>
<tr>
<td>Accept synonyms</td>
<td>Reveal ability to decipher meaning despite low literacy</td>
<td>Jongsma 1970,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Elley 1979</td>
</tr>
</tbody>
</table>
Table 4 Standard & Modified Cloze Test Scores by Years of Education

<table>
<thead>
<tr>
<th>Years of Education</th>
<th>6-8 (N=14)</th>
<th>9-10 (N=5)</th>
<th>11-12 (N=15)</th>
<th>Overall (N=34)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Standard</td>
<td>Modified</td>
<td>Standard</td>
<td>Modified</td>
</tr>
<tr>
<td># Correct</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>12</td>
<td>27</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>12</td>
<td>32</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>11</td>
<td>32</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>23</td>
<td>35</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>29</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>31</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>33</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>31</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>33</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean # correct</td>
<td>18.71</td>
<td>24.64</td>
<td>26.40</td>
<td>32.20</td>
</tr>
<tr>
<td>Mean % correct</td>
<td>37%</td>
<td>49%</td>
<td>53%</td>
<td>64%</td>
</tr>
</tbody>
</table>

Standard scores accept only exact replacements of words deleted from the text. Modified scores disregard spelling errors and accept synonyms. Standard cloze scores do not correlate with years of education. Adjusted R Squared = .099  P = .360. Modified scores are significantly correlated with years of education. Adjusted R Squared = .802  P = .000.
Table 5 Selected Participants' Comments

Well there are many people that do not know how to care for themselves during pregnancy, lots of people need information.

~ Mother of 4, 8 years education

...There are many things that I did not know until now, I have noticed. For example, right now I had a headache and I did not know why. I am pregnant. I read a paragraph that talked about a headache and it is okay.

~ Mother of 1, 10 years education

I never knew that in the second month a baby was all there. But, already their feet and hands are there.

~ Mother of 3, 12 years education

...I would do any sacrifice to leave something that would jeopardize my baby.

~ Mother of 4, 8 years of education

It gives us ideas of how to eat well, also it makes us react to the dangers of not caring well for yourself. ... It always reminds us that whatever problems, we should consult with the doctor because more than anything, one is not one life but two, and when we are in danger well, the baby is also in danger.

~ Mother of 4, 12 years education

I will be able to see what I should do, how I should take care of myself. I have to take care of myself so as to take better care of my baby so that he is healthy and in this book I will find that.

~ Mother of 2, 11 years education

It is well written in Spanish, but also it depends, for example, on the preparation that the persons have. There will be some persons that are not going to understand some words, but the book is good because it explains to one the development of the child, how it continues to develop in the abdomen... It helps you to think if one is doing something incorrect well try to stop it.

~ Mother of 4, 12 years education

[I will read it] because it is in Spanish ...and right now I have 3 children and I have never read about what is happening in my body.

~ Mother of 3, 12 years education

This book is good, explains lots of things about pregnancy. It also contains very important information about violence because in some cases one never knows who to trust ... but there are places, telephone lines that can help you. ... There are many persons that can find help inside of this [book], it is very good.

~ Mother of 2, 9 years education

[Some husbands would not like this book] because they would not like to know that we can get advice and that there is somebody to whom we can go when we are being abused or mistreated.

~ Mother of 3, 6 years of education

I had my last baby but I like to know lots so that I can teach other persons, perhaps my nieces.

~ Mother of 4, 8 years of education
Acknowledgements

The authors would like to thank the persons and institutions that provided the information and support necessary to carry out this study. CareOregon: Nancy Guenther, Christine Sovari, Abbey Stamp O’Connor, Mary Lou Hennrich, Carole Romm, Raquel Ramirez Mulato, Lupe Bermudez, Pamela Bermudez, Mark Moulton. University of Washington: Alicia Silva-Santisteban, Andie Baker, Jim Hughes. Translators: Benjamin Arenas, Patricia Zuluago and the City of Hartford Health Department Perinatal Outreach Program staff. We would also like to thank Annalisa Nixon.