

**Zeitschrift:** Studies in Communication Sciences : journal of the Swiss Association of Communication and Media Research

**Herausgeber:** Swiss Association of Communication and Media Research; Università della Svizzera italiana, Faculty of Communication Sciences

**Band:** 5 (2005)

**Heft:** 2

**Artikel:** Developing culturally and linguistically appropriate health education materials

**Autor:** Smith, Sandra A. / Gonzales, Virginia

**DOI:** <https://doi.org/10.5169/seals-790929>

### **Nutzungsbedingungen**

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

### **Conditions d'utilisation**

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

### **Terms of use**

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

**Download PDF:** 01.04.2026

**ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>**

SANDRA A. SMITH & VIRGINIA GONZALES\*

## DEVELOPING CULTURALLY AND LINGUISTICALLY APPROPRIATE HEALTH EDUCATION MATERIALS

This article reports results of pilot testing methods to evaluate the suitability of health education materials adapted from English source documents to serve non-English learners. We worked directly with monolingual Spanish speakers to adapt pregnancy health information to serve Spanish-speaking women in the United States. The project aimed to ensure that women of diverse cultures and varied literacy skills have access to essential prenatal care information according to practice guidelines for the United States and Canada, via a tested process for verifying the suitability of translated and tailored information. Specifically, the aim was to increase access to key messages linked to positive birth outcomes (Kogan et al. 1994, Davis & Akridge 1987, Libbus & Sable 1991). Cloze testing produced quantitative data and Reader Verification Interviews produced qualitative data. Previously validated modifications to the cloze test instrument and standard scoring procedure reveal ability to decipher meaning despite low literacy skills. These modifications produce more useful results when using the cloze test to evaluate the suitability of content and materials rather than the language proficiency of learners. Qualitative data support modified scoring. This project produced culturally and linguistically appropriate learning materials for Spanish-speakers and guidelines for adapting English health education materials to serve the needs of diverse populations.

*Keywords:* literacy, health literacy, cloze testing, tailoring materials, testing materials.

\*University of Washington, Seattle WA, sandras@u.washington.edu;  
gonza@u.washington.edu

## 1. Introduction

To ensure that women of diverse cultures and a wide range of literacy skills have access to essential pregnancy information according to prenatal care guidelines for the United States and Canada (US Public Health Service 1989), this project pilot tested methods for verifying the suitability of health education materials adapted from English to serve non-English learners. Specifically, the project aimed to increase access by Spanish-speakers in the United States to certain health behavior messages, which are linked to positive birth outcomes (Kogan et al. 1994; Davis & Akridge 1987; Libbus & Sable 1991).

## 2. Study Population

From January to August 1999, we recruited and interviewed study participants in Portland, Oregon in the northwest U S. Project staff randomly selected potential participants from the Spanish-speaking segment of women who obtained prenatal care through CareOregon, a not-for-profit managed care organization serving publicly insured residents of the state of Oregon. 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 (N = 35) ranged in age from 18 to 38, with an average age of 27 years. Most were married (83%) and not employed (74%). Most were born in Mexico (83%); 12% were born in the United States.

Nearly three quarters of the participants (71%) identified themselves as Latino or Hispanic and 10% identified themselves as *descendiente de mexicano* (descendants of Mexicans). Over half (57%) have lived in the United States 10 to 23 years. On average, participants have been US residents for 7 years. Half (51%) completed six to nine years of schooling and 46% completed 10-12 years. Average educational achievement was 9 years.

### 2.1. Test Materials

We selected source materials already rigorously reviewed for scientific accuracy and developed and tested with socioeconomically and culturally diverse populations 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 and insurance communities since 1989. They also have been used in Canada, Australia, New Zealand, Guatemala and Albania. The test materials, now in their sixth edition, are currently published under the title *Beginnings Pregnancy Guide* © 2005 ([www.beginningsguides.net](http://www.beginningsguides.net)). Readability ratings for the materials are fourth 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 nine years education reported extreme satisfaction with these materials (Smith 1998).

### 2.2. Translation

Since literal translation ignores differences among cultural groups and subtleties in the language, to achieve the goal of producing materials suitable for most Spanish-speakers in the US, we formed a partnership with the staff of the perinatal outreach program of the City of Hartford (Connecticut) 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 living on the east coast of the US. 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 the west coast city of Seattle, Washington incorporated the Mexican perspective and edited the text as necessary to fit the pages. The resulting Spanish language text rated a third grade reading level on the Spanish Fry Index (Gilliam et al. 1980, Crawford 1985). 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.

### 2.3. 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 sessions and had eight no-shows. Thirty-five participants completed the interview and cloze test - a suitable number for this type of formative research (Doak et al. 1996)

#### *2.4. Screening and Interviews*

A bilingual employee of CareOregon telephoned potential participants in the assigned random order. If a candidate 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, tested 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 session to be conducted at the most convenient of four clinic locations or in her home. We sent a confirmation letter to each participant who scheduled an appointment.

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

#### *2.5. Data Collection Procedures and Instruments*

We collected both qualitative and quantitative data to evaluate the suitability of the translated test materials. The Reader Verification and Revision Interview is a qualitative process to uncover specific content or format features that learners do not understand or do not accept, and to produce remedies (Doak et al. 1996). Cloze testing (Taylor 1953) also reveals specific language that readers may not comprehend and produces quantitative data in the form of scores. Collecting both qualitative and quantitative data provides a check and balance that helps assure accurate interpretation of findings.

One investigator (Gonzales) and a project staff person conducted a one-hour session in Spanish with each of the study participants. Each ses-

sion consisted of an introduction (5 minutes), cloze test (20-30 minutes), break and refreshments (5 minutes), Learner Verification and Revision Interview (20-30 minutes). The investigator explained the purpose of the study and confidentiality protections and obtained participants' permission to audio record interviews.

### *2.6. Learner Verification and Revision Interviews*

We selected for testing text and graphics that we suspected might be misinterpreted, irrelevant, or unacceptable. Each study participant viewed graphics and read sections of a prototype of the first booklet in a six-booklet series. The investigator conducted the Learner Verification and Revision Interview using an open-ended questionnaire designed to uncover specific content and format features that the participant did not understand or accept, and to produce solutions. Working with transcribed and translated manuscripts, we categorized responses to each question according to content and produced a summary.

### *2.7. The Cloze Test*

The cloze test is a self-administered paper and pencil test used to evaluate comprehension of educational material and language proficiency. The test method is validated in college students and the military (Foltz & Sullivan 1998). Weiss et al. (1992) and others used a cloze procedure to test adult patients' health literacy, defined as ability to understand health related information.

To create the instrument, every fifth word is deleted from a 250 word 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 knowledge the reader obtained from information surrounding the blanks. By requiring exact replacements of missing words, the cloze procedure also tests vocabulary and detects spelling and grammatical errors.

Rather than testing learners' language proficiency, we used the cloze test to evaluate the suitability of health related content and material as suggested by Doak et al (1996). Our intent was to a) confirm comprehension of the test material; b) to discover terms and concepts that readers found difficult or unacceptable; and c) to find in the "wrong" word replacements the specific language that learners expect or anticipate, and

therefore will read more easily and understand more readily and accurately. When testing information instead of learners, cloze test responses suggest ways to lighten the literacy demand of the information and so increase learners' comprehension.

A number of cloze test scoring procedures are documented in the literature (Elley 1979, Raggett et al. 1979). According to Busselman and Holcomb (1994) scores of 44% to 56% equate to 75% comprehension while scores  $\geq 57\%$  equate to 90% comprehension. Doak et al. (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
- $\geq 20$  correct ( $\geq 40\%$ ): Materials are useful. Assistance may be required for complete understanding
- $\geq 29$  ( $\geq 57\%$ ): Materials are easy to understand, suitable for independent learning.

Studies show that different forms of the cloze 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 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.

Cloze testing may induce anxiety, especially among those with negative or unsuccessful school experience, those embarrassed by their low literacy (Parikh et al. 1996), and those for whom immigration status is a concern. All these factors were common among our testers. Reducing anxiety is significant since Newcomer et al. (1999) showed that anxiety reduces cognitive ability. Rush and Klare (1978) validated a modification to reduce testers' frustration by using word-length blanks instead of uniform-length blanks. Busselman and Holcomb (1994) used this modification in cloze testing with a study population similar to ours. We used it as well. To further reduce participants' anxiety, the interviewers emphasized that the intent was to test the *materials*, not the participant, as suggested by Doak et al. (1998).

### 3. Results & Discussion

#### 3.1. Reader Verification and Revision Interviews

We evaluated five elements to determine the likelihood that the test materials would positively influence knowledge and health behaviors: attraction, comprehension, acceptability, self-efficacy and persuasion.

#### 3.2. 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. 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.

#### 3.3. Acceptability

To avoid unfamiliar words and to discover terms that would convey the same meaning to speakers of various Spanish dialects. Interviewers 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 nearly all 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.

We asked directly how we should talk about urination. Perhaps reflecting the more formal nature of Spanish, two-thirds of the women preferred the correct term (*orinar*); 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 pre-testing the translation with clients of the Hartford (Connecticut) Health Department, the east coast translation team noted that many clients preferred the less formal address *tu* (you) over the more formal and proper *usted*. This led the investigators to include a question about preference in the west coast interviews. The participants were equally divided in their preference. One-third (34%) preferred *usted* and one-third (31%) preferred *tu*. The remaining third of the testers said both forms were acceptable to them. Purchasers said they would not distribute materials using *tu*. We chose the more formal and respectful *usted*.

### 3.4. *Comprehension, Self-efficacy and Persuasion*

Comprehension is critical for persons with low literacy since they can access fewer information resources than highly literate learners. The interviewer asked participants to read a section titled “Keys to a Healthy Baby” and then name “*healthy things to do during pregnancy*”. All successfully named several 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 – participants’ feelings of confidence that they can do what is required to have a healthy baby (Bandura 1986) – the interviewer asked which key message seemed most difficult for the participant 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 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 advice to gain weight during pregnancy have a significantly reduced chance of delivering a low birth weight infant (Kogan et al. 1994).

After reading a list titled *Warning Signs*, 100% of the testers named on average 3 or 4 reasons to call the doctor, demonstrating good comprehension. 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; 14% said they would call or visit the hospital or clinic. If the women suspected a problem but were not sure that it warranted a call to the doctor, 86% said correctly that they would call or visit the doctor anyway. This shows that the test materials are likely to increase women's awareness of warning signs and persuade them to trigger timely intervention.

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 may recall only the last few words of a sentence.

Interviewers 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.

No participant was offended by anything in the test materials, and all said they would like to receive the other five booklets in the series. Participants' comments shown in Table I below 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.

*Table I: Mothers use information to make healthful changes*

<i>I will read it because it is in Spanish ...and right now I have 3 children and I have never known what is happening in my body.</i>	Mother of 3, 12 years education
<i>I would do any sacrifice to leave something that would jeopardize my baby.</i>	Mother of 4, 8 years education
<i>I never knew that in the second month a baby was all there. But already their feet and hands are there.</i>	Mother of 3, 12 years education
<i>[Some men] 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.</i>	Mother of 3, 6 years education
<i>I had my last baby, but I like to know lots so that I can teach other persons, perhaps my nieces.</i>	Mother of 4, 8 years education
<i>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.</i>	Mother of 2, 11 years education

While the interview data showed high levels of comprehension, acceptability and persuasiveness, the cloze scores suggested low comprehension and indicated the materials were too difficult for about one third of testers (29%) to use independently. Seeking to explain the apparent contradiction between our qualitative and quantitative results, we expanded our review of the scientific literature on cloze testing. We found relevant reported difficulties with the scoring procedure along with remedies, which we employed. Standard cloze test scoring accepts only exact replacements of words deleted from the sample text. We were concerned only with comprehension (decoding the meaning) and not with other literacy skills that standard scoring is designed to demonstrate, such as spelling and grammar. Jongasma (1970) documented higher reliability of

cloze test scores with synonyms counted; as did Holcomb (1978) and Elley (1979). Elley also validated scores that disregarded misspellings. Accepting these researchers' modifications, we calculated a second modified cloze score that accepted synonyms and disregarded spelling and grammatical errors.

While Elley (1979) reported that accepting synonyms and disregarding spelling errors increased high school students' number of correct entries by only 3 or 4 in scientific material and by only 1 or 2 in prose, in our sample of disadvantaged mothers long out of school with average 9 years of education, the modification resulted in a much larger typical gain of 6 or 7 ranging up to 11 correct entries. These gains demonstrated that participants were often able to decipher the meaning although their literacy 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.

At the same time, the synonyms considered "wrong" in the standard scoring procedure identified terms more familiar and meaningful to the readers. 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' obstetriz*. 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 testers understood the meaning and cued us to change the wording to participants' most common response.

Modified cloze test scores reflected the quantitative data more closely than standard scores. Modified scores showed that 83% of testers found the materials easy to use for independent learning. Table II below illustrates that with synonyms accepted, 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%. The overall average score increased from 48% to 60%. These modified averages are well above the 57% indicator of full independent comprehension.

Table II: Mean Standard & Modified Cloze Scores by Years in School

Years in School	6-8 (N=14)		9-10 (N=15)		11-12 (N=15)		Overall (N=34)	
	Standard	Modified	Standard	Modified	Standard	Modified	Standard	Modified
Mean # correct	18.71	24.64	26.40	32.20	27.53	34.47	23.74	31.12
Mean % correct	37%	49%	53%	64%	55%	69%	48%	60%

Standard scores accept only exact replacements of words deleted from the text. Modified scores disregard spelling errors and accept synonyms.

These findings indicate that the test materials are suitable for independent learning for 83% of women in this population of publicly insured disadvantaged 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 *independently*. About 17% need assistance to fully understand and apply the information. One interpretation is that this group lacks sufficient literacy skills to learn *independently* from materials written at a 3rd grade level. Another interpretation is that the suitability of the material for this audience can be improved. Testers' responses in the interviews and cloze tests suggested specific text revisions to make both the source material and the translated material easier to read and apply for all readers.

### 3.5. Demographic Factors & Cloze Test Scores

Using SPSS for Windows Version 9.0, we performed univariate and multivariate analysis to determine effects of demographic factors on standard and modified cloze test scores.

### 3.6. Years of Education

Since the cloze test is a school-based measure so that those with more formal education are likely to be more experienced and more comfortable with this type of testing; since years of formal education is an indicator of literacy level in populations (Doak et al. 1996); and since individuals' cloze test scores on health related material have been positively associated with educational achievement (Rutledge and Donaldson 1998), we expected scores for the study population to reflect years of education. 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,  $p = .000$ ). This difference is most likely due to the fact that modified scores consider only comprehension by accepting synonyms and disregarding spelling and grammatical errors, which we considered to be irrelevant since we were testing the material, not the learners. With more education and life experience, a person may increase comprehension (ability to decode meaning) while not improving spelling or grammar.

As in previous studies (Doak et al. 1996; Rutledge & Donaldson 1998), in this project the strength of the correlation between educational achievement and reading comprehension scores varied widely among individuals. Years of education did not reliably predict individual scores. Notably, both the highest score – 35 (42 modified) correct – and the lowest score – 6 (8 modified) correct – were achieved by women who each obtained 9 years education in Mexico.

For those with 6 to 8 years education, the range of standard scores was 6 to 32 of 50 possible correct; their modified scores ranged from 7 to 41 correct. For those with 11-12 years education, the range of standard scores was 13 to 34 correct; their modified scores ranged from 18 to 39 correct. Table II shows standard and modified cloze scores by years of education.

### *3.7. Years in the United States*

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 showed a significant correlation with years in the country (adjusted R squared = .513,  $p = .041$ ). This finding further suggests that modified scoring may more accurately reflect ability to decipher meaning despite low literacy skills.

Notably, although the study participants have resided in the United States on average 7 years and up to 23 years, all but one remain monolingual. Persons not proficient in English face substantial communication barriers at every level of the health care system (US Department of Health & Human Services 1995). 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 speaking residents and citizens.

### 3.8. Number of Children

We reasoned that “experienced” mothers would have more previous knowledge of pregnancy and prenatal care so that comprehension and 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 modified scoring may produce scores that more accurately reflect previous knowledge and ability to decipher meaning despite low literacy skills. Table III below shows demographic factors explaining the differences in scores.

*Table III: Demographics & Differences in Cloze Scores*

Cloze Score →	Standard	Modified Adjusted R Squared
Years of Education	No correlation	.802 (p=.000)
Years in US	No correlation	.513... (p=.041)
Parity	No correlation	.515.... (p=.041)

## 4. Summary and Conclusions

This project pilot tested processes and instruments to evaluate the suitability of health education material translated and adapted from a source document in another language. Specifically, we tested prenatal education materials adapted from English to serve Spanish-speaking women in the United States. 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.

Since the two evaluation methods initially produced contradictory results, we adopted validated modifications to the cloze test scoring procedure to see if modified scores would more closely reflect interview findings and observations. We analyzed effects of demographic factors on

both standard and modified cloze test scores and 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 supported modified scoring. This suggests that the modifications to the cloze test scoring procedure take into account previous knowledge and reveal ability to decipher meaning despite low literacy skills. The standard scoring procedure, which requires exact replacement words, seems to underestimate entry-level knowledge and ability to comprehend health information that is applicable and important to the learner.

Accepting the modified scoring, the test materials were demonstrated suitable for 83% of women in this population of publicly insured disadvantaged 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 *independently*. About 17% did not could not use the materials before improvements suggested by the process were implemented.

In this pilot project, the methods 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 suitable for independent learning. Future research should include retesting to judge the effects of improvements and rule out unintended effects.

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 population.

#### *4.1. Guidelines for Adapting Materials to Serve a Diverse 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 populations.

#### *4.2. Materials*

Select source materials that are demonstrated suitable and effective for their intended audience. Consider content and design factors that affect comprehension, as well as reading level as well as. SAM, the Suitability Assessment of Materials (Doak et al. 1996) is a validated scored instrument useful in evaluating source materials.

Recognize differences among cultural groups and subtleties in the language. Avoid literal 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.

#### *4.3. Project Management*

Ensure that the budget allows for the extra time, resources and expertise that a bilingual, cross-cultural project requires. This project was allocated \$50,000 US. Staff and collaborators absorbed the cost of extra travel and time required for recruitment and interviewing. The budget did not allow for retesting to verify that revisions produced the intended improvements without unanticipated effects.

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.

#### *4.4. Instruments*

The Reader Verification and Revision Interview procedure reveals previously held beliefs and previous knowledge as well as concepts that do not come across in translation. This interview process uncovers and suggests

remedies to elements that may be misinterpreted. Asking the meaning of specific terms is useful.

The cloze test procedure 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. When testing materials, not the learner, the instrument should be modified to reduce anxiety and scoring should be modified to allow misspellings and synonyms. These modifications reveal testers' ability to decipher meaning despite low literacy skills and differentiate 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.

*Bottom Line:* Learners have the last word.

*Acknowledgements:* This project was funded by the US Public Health Service Agency for Health Care Research & Quality. We further acknowledge CareOregon and University of Washington Center for Health Education & Research for their cooperation and assistance.

## References

- BANDURA, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*, Englewood Cliffs, NJ: Prentice Hall.
- BUSSELMAN, K.M. & HOLCOMB, C.A. (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. (1987). The effects of promoting intrauterine attachment in primiparas on post delivery attachment. *Journal of Obstetric, Gynecologic, and Neonatal Nursing* 16/6: 430-7.
- DOAK, C.; DOAK, L. & ROOT, J. (1996). *Teaching Patients with Low Literacy Skills*, Philadelphia: Lippincott.
- DOAK, C. et al. (1998). Improving Comprehension for Cancer Patients with Low Literacy Skills: Strategies for Clinicians. *Cancer Journal for Clinician* 48:151-162.
- ELLEY, W. (1979). A close look at the cloze test. *Research Information for Teachers, Special Issue*.
- FOLTZ, A. & SULLIVAN, J. (1998). Get Real: Clinical testing of patients' reading abilities. *Cancer Nursing* 21/3: 162-166.
- GILLIAM, B.; PENA, S.C. & MOUNTAIN, L. (1980). The Fry graph applied to Spanish

- readability. *Reading Teacher* 33/4: 426-30 .
- CRAWFORD, A.N. (1985). Formula y graphico para determinar la comprensibilidad de textos del nivel primario en castellano. *Lectura y Vida* 6/4: 18-24.
- HOLCOMB, C.A. (1978). The Cloze Procedure, *Health Education* 9/6: 8-10.
- JONGSMA, E.A. (1970). The Cloze Procedure: A survey of the research. *Occasional Papers in Reading*. Bloomington: Indiana University School of Education.
- KOGAN, M.D. et al. (1994). Relation of the Content of Prenatal Care to the Risk of Low Birth Weight. *Journal of the American Medical Association* 271/17: 1340-1345.
- LIBBUS, M.K. & SABLE, M.R. (1991). Prenatal Education in a High Risk Population: The Effect on Birth Outcomes. *Birth* 18/2: 78-84.
- NEWCOMER, J.W. et al. (1999). Decreased memory performance in humans induced by stress-level cortisol treatment. *Archives of General Psychiatry* 56: 527-533.
- PARIKH, N.S. et al. (1996). Shame and health literacy, the unspoken connection. *Patient Education and Counseling* 27: 33-39.
- RAGGETT, M.J.; TUTT, C. & RAGGETT, P. (1979). Assessment and Testing of Reading: Problems and Practices, London: Biddles of Guildford.
- RUSH, R.J. & KLARE, G.R. (1978). Reopening the cloze blank issue. *Journal of Reading Behavior* 10: 208-10.
- RUTLEDGE, D.N. & DONALDSON, N.E. (1998). Improving Readability of Print Materials in Patient Care and Health Services. *The Online Journal of Clinical Innovations* 1/3: 1-27.
- SMITH, S. (1998). Information Giving: Effects on Birth Outcomes and Patient Satisfaction. *International Electronic Journal of Health Education* 1/3:135-145. Retrieved on July 1998 from [www.beginningsguides.net/infogiving](http://www.beginningsguides.net/infogiving).
- TAYLOR, W.S. (1953). CLOZE procedure: A new test for measuring readability. *Journalism Quarterly* 30: 415-433.
- US Department of Health and Human Services, Office for Civil Rights. (1995). Limited English Proficiency as a Barrier to Health & Social Services, Washington, DC: Macro International.
- US Department of Health and Human Services. (1993). Draft Limited English Proficiency Regulation, 45. Code of Federal Regulations, Part 80. (on file with National Health Law Program).
- US Public Health Service. (1989). Caring for Our Future: The Content of Prenatal Care. Washington, DC: Department of Health and Human Services.
- WEISS, B.D. et al. (1992). Health Status of Illiterate Adults: relations between literacy and health status among persons with low literacy skills. *Journal of the American Board of Family Practitioners* 5: 257-264.
- WHEILDON, C. (1995). Type and Layout: How typography and design can get your message across or get in the way, Berkeley, CA: Strathmoor Press.