

3. READABILITY OF CAHPS® 2.0 CHILD AND ADULT CORE SURVEYS

Background

In recent years, the emergence of managed care has prompted interest in collecting survey information from health care consumers. Many public and private purchasers of care either already administer patient surveys to their beneficiaries or plan to in the near future. However, the growing diversity of the U.S. population poses major challenges for developing such survey instruments. First, the cultural and linguistic diversity of many beneficiary groups requires that surveys be appropriately translated into various languages and adapted for different groups. Second, because patient surveys are often self-administered, attention must be given to survey readability.

Research studies from many sources, including national literacy data, tell us that a large share of U.S. adults can only read at very basic levels. This problem is particularly striking among Medicaid beneficiaries. According to the 1993 National Adult literacy survey (Kirsch, Jungeblut, Jenkins & Kolstad, 1993), 75% of welfare recipients read at or below the eighth grade level and 50% read at or below the fifth grade level.

Moreover, low reading skills may be more concentrated among certain Medicaid beneficiary sub-groups more than others. For instance, immigrants and refugees from less-developed countries may be more likely than U.S.-born Medicaid beneficiaries to have low educational attainment and, as a result, low reading skills. Among recent Central American immigrants and refugees entering the U.S. from El Salvador and Guatemala, nearly 80% reported less than a high school

education (Lopez, 1996). Among foreign-born Hispanics living in the Los Angeles region, 10% report no schooling, 38% reported elementary school only, and 21% reported some high school (Cheng & Yang, 1996).

The mismatch between an intended respondent's reading ability and the survey instrument may have important implications for the validity of patient satisfaction research, particularly for self-administered surveys. Some of the consequences of this mismatch may include low response rates, especially in vulnerable populations, and unreliable responses because of poor item comprehension.

This study assesses the readability of the English and Spanish versions of the Consumer Assessments of Health Plans Study (CAHPS®) 2.0 adult and child core surveys. The linguistic and cultural adaptation of these surveys is discussed in a separate paper (Weech-Maldonado, Weidmer, Morales & Hays, 1999).

The CAHPS® Surveys

CAHPS® is a 5-year initiative that aims to produce a set of standardized survey instruments that can be used to collect reliable information from health plan enrollees about the care they have received and their experiences with their health plan. The results of the surveys are turned into reports that provide decision support to consumers selecting a health plan.

To date, several instruments have been developed as part of this study, each targeting a specific population served by health plans throughout the U.S. CAHPS® has also developed surveys for children, designed for a proxy respondent. Although variations exist between the different versions of these instruments depending on the target

population and the age of the respondent, a core set of survey questions is common to all versions of the survey. Five specific domains of care (getting needed care, getting care quickly, communication with providers, office staff courtesy and respect, health plan customer service) and global ratings (care overall, personal doctor or nurse, specialist care, health plan) are assessed in the CAHPS® 2.0 surveys.

The CAHPS® investigators recognized the need to translate its instruments into other languages. Indeed, the CAHPS® survey instruments were translated into Spanish (Weidmer, Brown & Garcia, 1999) because many participating health plans are located in states that have large numbers of Spanish speakers, including Texas, California, New Jersey and Florida. Hence, we evaluate the Spanish versions of the adult and child core surveys along with the English surveys.

Assessing Readability

Two major approaches are available for assessing the readability of documents - measurement and prediction. Measuring readability, by judgment or comprehension tests, involves using readers. Readability by judgment is usually obtained by asking literacy experts to determine the readability level of a document based on their experience or on use of an algorithm. Readability by comprehension test is obtained by administering a reading comprehension test based on the written material to readers of known ability. A test score criterion is chosen that defines comprehension of the material. When some proportion of readers of similar ability achieve that score, the reading ability of the test takers corresponds to the readability level of the document.

In the second approach, mathematical formulas predict the readability of a document. Unlike judgments or comprehension tests, readability formulas do not rely on readers to establish the readability level of written materials. Because no measurements are made, readability formulas are strictly prediction tools.

The selection of readability technique depends upon the factors of time, availability of subjects, level of resources available to conduct the assessment, and the degree of accuracy required in assessing the materials for the target groups (Klare, 1974). Predicting readability by formulas does not involve readers and is therefore much less expensive, but it only provides an approximate indication of the readability of a document. Measurements obtained by tests and judgments by experts require greater resources but provide more accurate assessment of readability. We chose to use the former approach for this study for two reasons. First, prior research had addressed the readability of the CAHPS® surveys through cognitive interviews (Harris-Kojetin, Fowler, Brown, Schnaier & Sweeny, 1999) and expert judgments (Brown, Nederend, Hays, Short & Farley, 1999). Second, available resources constrained us to using readability formulas.

To identify appropriate readability formulas for our study, we conducted a literature search. Our goal was to identify formulas appropriate for survey instruments in Spanish and English. Although we found references to numerous readability formulas, we did not identify any formulas appropriate for evaluating survey instruments in English or Spanish. The principal problem with applying readability formulas to survey instruments is that the formulas become unreliable when applied to passages of fewer than 100 words (Fry, 1990). Because the

CAHPS® surveys are composed of multiple closed-ended questions followed by a set of response options, passages of less than 100 words are common. Furthermore, the vast majority of formulas we identified were appropriate for English written materials but not Spanish.

Most readability formulas typically use two factors in their calculations: a sentence or syntactic factor and a word or semantic factor (Rush, 1985). Formulas using these two factors include the Fry Readability Graph (Fry, 1965), Dale-Chall (Dale & Chall, 1948), Fog (Gunning, 1968), Flesch (Flesch, 1948), and Flesch-Kincaid (Kincaid, Fishburne, Rodgers & Chissom, 1975). The SMOG (McLaughlin, 1969) is an exception because it has only a syntactic factor. The syntactic factor frequently estimates the grammatical complexity of the writing by using sentence length. The semantic factor purports to measure the degree of difficulty of the vocabulary in a piece of writing. Readability formulas usually estimate semantic load either with a measure of word length such as number of syllables or with a count of unusual words. Thus the assumption that word and sentence length are reasonable correlates of syntactic complexity and semantic load underlies reading formulas (Rush, 1985).

Readability formulas are typically validated against performance criterion passages of varying but known levels of difficulty. Two common sources of criterion passages are the McCall-Crabbs Standard Test Lessons in Reading (McCall & Crabbs, 1961) and the Gates-MacGinite reading tests (Gates & MacGinite, 1965). The validity of a particular readability formula is determined by how accurately it predicts the grade level of a criterion passage. In addition, the validity of more recent formulas is established in part through correlation with older formulas.

In addition to using a readability formula, some investigators have chosen to describe readability using a variety of counts of syntactic and semantic factors (Leadbetter, 1990). Fry recommends the use of word counts and sentence length to assess the readability of passages having fewer than 100 words (Fry, 1990). Because readability formulas were not originally intended for survey instruments, we have supplemented the readability formula results with counts of a variety of syntactic and semantic factors (see Table 1, page 73).

Adapting Survey Instruments for Readability Assessments with Formulas

Using readability formulas to assess the CAHPS® surveys required us to exclude the question response scales, leaving only the instructions, question preambles and the survey questions themselves. The question response scales were deleted from the text of the surveys because they do not have a sentence structure, which readability formulas assume.¹

Fry Readability Graph

The Fry Readability Graph (FRG) is the principal readability assessment tool used in this study because it has been validated for Spanish and English language documents. Like most readability formulas, the FRG has syntactic and semantic factors - sentence length and syllables. To implement the FRG, one first randomly selects three sample passages of exactly 100 words - from the beginning, middle, and

¹ Other researchers have turned response options into sentences and included them in their readability analysis (Lewis, Merz, Hays, and Nicholas, 1995).

end of the source document. (Our source documents consisted of the CAHPS® surveys, stripped of all response options.) After the total number of sentences and syllables for each of the 100-word passages has been recorded, the average number of sentences and syllables is computed. The resulting figures are plotted on a graph and the resulting coordinate point is associated with an established grade level designation. An illustration of the FRG is shown in Figure 1. The FRG is appropriate for assessing materials from the first grade through the college level (Fry, 1969, 1977).

The FRG is one of the few readability assessment tools that is adapted for Spanish language documents (Gilliam, Peña & Moutain, 1980). Spanish language application of the FRG is similar to its English language application, with the exception of syllable counting. Because of differences in the structure of words in the two languages, the syllable counts for 100-word passages in Spanish tend to be much higher than for the same passage in English. To correct for this discrepancy, 67 is subtracted from the total syllable count for each 100-word passage in Spanish (Gilliam et al., 1980).

The comparability of the FRG applied to Spanish language documents (with the adaptation) and English language documents has been assessed. Using Spanish primary textbooks, the readability level of the FRG and the publisher's grade level were compared. In 10 of 12 cases, the FRG level grade level and publisher's grade level were the same (Gilliam et al., 1980). Unfortunately, a similar comparability study has not been conducted using the FRG for documents at higher reading levels.

FRASE Graph

The FRASE graph is a readability assessment tool specifically developed for Spanish written materials (Vari-Cartier, 1981). The FRASE graph addresses two limitations of the FRG by increasing the syllable count range beyond 182 per 100-words and altering the readability designations from grade levels to the reading difficulty designations used in English as a Second Language instruction (Beginning, Intermediate, Advanced Intermediate, and Advanced).

The FRASE graph is derived from the FRG, also basing its readability assessment on a syllable and sentence count. However, the FRASE graph uses five 100-word samples rather than three.

The FRASE graph has been extensively validated using subjective teacher judgments, Spaulding formula scores, cloze test scores, and informal multiple-choice test scores. Correlation coefficients between the FRASE graph readability designations and the alternative readability estimates ranged from 0.91 to 0.97, indicating that the FRASE graph is equivalent to other established methods for estimating readability (Vari-Cartier, 1981).

Fog Index

The Fog Index, which uses as few as 100 successive words to determine both sentence length and the number of words with three or more syllables, was developed by Gunning (1968). The counts are then substituted into a formula² and the reading difficulty is calculated according to formal grade level in school. For longer written works, the author recommends selecting several 100 word samples from various parts of the material averaging the results to determine the reading

² FOG Readability formula: Grade Level = $0.4 * (\text{average sentence length} + \text{percentage of words with 3 or more syllables})$.

level. This formula is appropriate for assessing materials from the fourth grade through the college level.

The Fog Index has not been adapted for Spanish language materials.

SMOG Grading Formula

The SMOG grading formula is based solely on syllables. It was developed by G. Harry McLaughlin as a fast and accurate test of readability (McLaughlin, 1969). The SMOG Grading Formula estimates the grade level of a document by counting the number of polysyllabic words (words with 3 or more syllables) in three chains of 10 consecutive sentences taken from the beginning, middle, and end of the document being assessed.

An advantage of the SMOG is that the standard error of the readability prediction has been estimated (SE=1.5 grades) based on validation studies using the McCall-Crabbs passages. A standard error of 1.5 grade levels means that the material being tested will be fully comprehended³, by 68% of its readers who have reached a reading skill level within 1.5 grades of the SMOG score.

The SMOG grading formula has been adopted by the National Cancer Institute as the preferred method for assessing the readability of cancer communications after a comprehensive review of advantages and disadvantages, including how well alternative formulas predict readability (Romano, 1979).

³ The reading ability, indicated by the grade placement score, needed to answer 100% of test questions on the McCall-Crabbs passage for that grade level (Klare, 1974).

The SMOG grading formula has not been adapted for Spanish language materials.

Flesch Reading Ease Score

The Flesch Reading Ease Score is one of the most widely used readability assessment formulas. Rudolf Flesch published his first reading formula in 1945, based on the number of affixes, the average sentence length, and the number of personal references. He subsequently introduced the Reading Ease formula, which is based on number of syllables per 100 words and average number of words per sentence. When applied to a document, the Flesch Reading Ease formula results in a number ranging from 0 to 100. The lower the score, the more difficult the material is to read and comprehend. The Flesch Reading Ease Score has been validated against the McCall-Crabbs passages (Klare, 1974).

Studies have shown that scores of 90-100 characterize most comic books, scores of 60-90 characterize articles from the popular press (e.g., *Better Homes and Gardens*, *Newsweek*), and scores of 20-30 characterize reports from medical journals (e.g., *Journal of the American Medical Association*, *New England Journal of Medicine*) (Morrow, 1980).

In the computer adaptation of the Flesch Reading Ease formula, the syllable count is replaced by a vowel count, something computers can do more easily. Research by Coke & Rothkopf (1970) has showed that counting vowels provided very similar estimates as counting syllables.

The Flesch Reading Ease formula has not been adapted for Spanish language materials.

Findings

English Language Adult and Child CAHPS® Surveys

Table 2 (page 74) shows the readability formula and word and sentence difficulty results for the CAHPS® 2.0 adult and child core English language surveys. The average number of sentences and the average number of syllables are the main indicators of syntactic and semantic complexity used in all readability formulas except the SMOG. The average number of sentences per 100-word sample was 5.1 for the adult survey and 7.9 for the child survey. The average number of syllables per sentence per 100-word sample was 134.0 for the adult survey and 124.3 for the child surveys. In general, lower readability (less difficult) is assigned to written material that has shorter sentences and fewer syllables. The lower average number of sentences and higher average number of syllables in the adult survey may explain why the Flesch Reading Ease score for the adult survey is lower than that for the child survey (a lower score indicates more difficult text).

The FRG scores can be verified by plotting the average number of sentences and average number of syllables on Figure 1. The FRG results show that for both the adult and child surveys, a 7th grade reading level is required for comprehension.

Applying the FOG Index to the adult and child surveys resulted in similar but not identical results. With the FOG Index, readability levels of 8th grade for the adult survey and 6th grade for the child survey were obtained. These results are consistent with the higher Flesch score obtained for the child survey than adult survey.

Recall that the SMOG Readability formula relies exclusively on counts of polysyllabic words found in three strings of 10 consecutive sentences selected randomly from the written material. Results from analyses using the SMOG are in agreement with results using the Fry graph that a 7th grade reading level is required for comprehension of both the adult and child surveys.

In Table 2, we also show the results of the readability formulas applied to a children's story (Kayner, 1999) and an article from national newspaper (*New York Times*, August 23, 1999: A1, A23). The FRG, Flesch Reading Ease score, FOG Index, and SMOG consistently rated the newspaper article at a higher level than either survey. The results of these analyses place the children's story at a reading level near that of both surveys.

Table 2 also shows the results of counts of syntactic and semantic components of the surveys and other materials. The sentence complexity counts (words per sentence, syllables per sentence, and characters per sentence) indicate that the adult survey had greater sentence complexity than the child survey and the *Cricket* reader, and less sentence complexity than the newspaper article. The counts of semantic factors (1-syllable words, words with 2 or more syllables, number of characters per word, and number of syllables per word) are less easy to interpret. The newspaper had a greater average number of characters and syllables per word than either survey or the *Cricket* reader, indicating a greater use of longer words. The newspaper had a lower average number of 1-syllable words and a greater average of words with 2 or more syllables, also indicating a greater use of longer words.

Spanish Language Adult and Child CAHPS® Surveys

Table 3 (page 76) shows the readability formula and word and sentence difficulty results for the adult and child Spanish language surveys. The average number of sentences per 100-word sample was 6.8 for the adult survey and 4.4 for the child survey. The average number of syllables⁴ per sentence per 100-word sample was 202.0 for the adult survey and 194.3 for the child surveys. Although the adult survey has more sentences and more syllables than the child survey, the results of the FRG indicate a 7th grade reading level for both.

The FRASE graph uses a similar method to the FRG to assess the readability of Spanish language materials. The FRASE graph results indicate that both the adult and child surveys require an *intermediate level* of reading skill to be fully comprehended. While the FRASE graph was intended to gauge the difficulty of materials used to teach Spanish as a second language, these results provide a useful indication of the readability level of the surveys. Furthermore, they provide a means of assessing the comparability of the child and adult survey readability levels.

Table 3 also shows the results of the readability formulas applied to an article from a Los Angeles Spanish language newspaper entitled "Resultado mixto en reduccion de clases" (*La Opinion*, June 24, 1999: A1) and a beginning reader, *Aventuras* (Freeman & Freeman, 1997). Both the FRG and FRASE graphs rate the readability of the surveys lower than the newspaper and higher than the reader.

⁴ Unadjusted for the greater average number of syllables in Spanish language materials than in English language materials.

The syntactic counts (number of words, number of syllables, and number of characters) indicate that on average, the child survey sample had longer sentences than the adult survey. The semantic counts (number of characters and number of syllables per word) for both surveys were similar. The semantic counts of 1- and 2-syllable words were dropped from this analysis because the higher number of syllables in Spanish language materials makes them unreliable indicators of vocabulary complexity.

Discussion

The results of this study suggest that the CAHPS® 2.0 adult and child core surveys require a 7th grade reading level for adequate comprehension. The SMOG and Fry graphs both resulted in a 7th grade level readability assessment for the English language adult and child surveys. However, the FOG Index and the Flesch Reading Ease Score indicate that the adult survey may have a higher readability requirement than the child survey. This discrepancy may be due to greater sensitivity of the FOG Index and Flesch formulas to differences in number of sentence and/or number of syllables between the adult and child surveys than either the SMOG or FRG. While the FOG Index suggests that the magnitude of the difference between the adult and child surveys may be as great as 2 grade levels, it is difficult to determine the significance of the difference between Flesch scores of 71.3 and 89.6, since these scores are not tied to specific grade levels.

This study also shows that the English and Spanish versions of the CAHPS® surveys have comparable readability levels. Based on the Fry graph, both the English and Spanish adult and child versions for

the core CAHPS® surveys have 7th grade readability levels. The similarity of the readability levels provides support for the success of the translation from English to Spanish.

Although the 7th grade reading level may be appropriate for commercially insured populations, it may be too high for Medicaid populations. According to the National Adult literacy survey (Kirsch et al., 1993), as many as 75% of welfare recipients read at or below the eighth grade level and 50% read at or below the fifth grade level. This suggests that the reading level required by the CAHPS® core surveys for full comprehension exceed the reading ability of more than 50% of welfare recipients. When one considers particular Medicaid beneficiary subgroups, the mismatch may be even greater.

A recent Public Policy Institute of California study reported that 42% of California Medicaid beneficiaries had less than a high school education (MaCurdy & O'Brien-Strain, 1997). Among recent immigrant Medicaid beneficiaries, 54% had less than a high school education; among Hispanic immigrant Medicaid beneficiaries who had arrived in the U.S. before 1985, 71% had less than a high school education. Since self-reported educational attainment tends to overstate literacy, the problem of low literacy and illiteracy among these groups is likely to be dramatic.

Poor comprehension of survey questions among those responding to patient surveys may also lead to unreliable results. For instance, adults with low literacy skills may not comprehend the term "health insurance plan." Indeed, cognitive interviews suggested that Medicaid beneficiaries frequently rated their overall care when asked to rate their health plan (Brown, 1996; Brown et al., 1999). Cognitive

interviews also found that Medicaid beneficiaries had trouble understanding the concept of a primary care provider or regular doctor and had trouble differentiating between a health plan and Medicaid (Brown, 1996; Brown et al., 1999).

Limitations of Readability Formulas

It is widely acknowledged that reading is an interactive process that occurs between the text and the reader. In fact, research shows that readers use experiences, knowledge, and information processing skills to comprehend text (Johnston, 1983).

Readability formulas, being strictly text-based, do not address the interactive nature of the reading process. Most reading formulas, including those used in this study, employ syntactic and semantic factors and do not directly address factors related to communicating meaning. For instance, readability formulas do not distinguish between written discourse and nonsensical combinations of words (Dreyer, 1984). Moreover, formulas cannot assess other critical factors such as the reader's interest, experience, knowledge or motivation, all of which may influence the reader's ability to comprehend the cognitive task asked by a survey (Duffy, 1985). Other factors related to readability and not assessed by a readability formula include typographical and temporal factors (e.g., time allotted to complete the reading task).

According to a recent paper on communicating with Medicaid beneficiaries, producing readable health materials requires thinking carefully about the audience to assess whether the intended respondents have the information with which to respond to the kinds of questions the survey asks (Hibbard, et al., 1997). It means organizing the material covered by the survey to make the survey easier to respond to,

and eliminating extra material that can overflow a page and overwhelm the survey respondent. It also means formatting a survey so that the instructions are simple to follow and using 12- to 14-point serif type, ample margins and headers to aid in organization. Finally, the overall content and design of the survey must be friendly, appealing and culturally appropriate to gain respondents' attention and increase their comprehension of important messages (Root & Stableford, 1999).

Many of the domains mentioned in the paragraph directly above were addressed during the development of the CAHPS® surveys. Cognitive interviews were used to identify items or terms that were difficult to comprehend, questions that were misinterpreted, and response options that that were inappropriate for the question or failed to capture the respondents' true experience (Harris-Kojetin et al., 1999). Literacy experts were consulted to improve readability of the survey (Brown et al., 1999). And careful translation procedures were followed to ensure the comparability of the English and Spanish versions of the surveys (Weidmer et al., 1999). These efforts provide additional evidence of the overall quality of the CAHPS® surveys.

This study is not intended to provide the definitive assessment of the readability of the CAHPS surveys. Rather, it aims to provide an additional rough gauge of their readability. Incidentally, a readability assessment by two literacy experts placed the readability level of the CAHPS surveys between the 6th and 7th grades (Julie Brown, personal communication, August 20, 1999).

Conclusions

Although the current readability level of the CAHPS® surveys may be appropriate for commercially insured populations, lower readability

is desirable for those who are publicly insured. As many as 50% of welfare recipients may fail to respond to the CAHPS® surveys because of a mismatch between the readability level of the surveys and the reading level of the intended respondents.

This situation may be exacerbated for certain subgroups of Medicaid beneficiaries, such as immigrants and refugees from less developed countries. According to research, non-English speaking patients and patients with low literacy skills face the greatest threat of receiving poor quality of care (Baker, Parker, Williams, Pitkin, Parikh, et al., 1996; Morales, Cunningham, Brown, Lui & Hays, 1999). Paradoxically, patients with low literacy skills also face the greatest barriers to responding to self-administered quality assessment tools such as the CAHPS® surveys.

Lowering the readability of the CAHPS® surveys, however, may be difficult. For reports about the CAHPS® surveys to help consumers make an informed choice about their health plan, the surveys need to collect information on a range of complex topics that require respondents to be familiar with concepts and vocabulary unique to health care. Shortening the survey and simplifying the vocabulary too much may cause the level of information gleaned from the CAHPS® surveys to fall, defeating the original purpose of CAHPS®.

Finding a balance between collecting important information and maintaining a reasonable level of survey readability will be an important consideration for researchers as future versions of the CAHPS® surveys are developed.

References

- Baker, D., Parker, R., Williams, M., Pitkin, K., Parikh, N., Coates, W. & Imara, M. (1996). The health care experiences of patients with low literacy. *Archives of Family Medicine*, 5, 329-334.
- Brown, J., Naderend, S., Hays, R., Short, P. & Farley, D. (1999). Special issues in assessing care of medicaid recipients. *Medical Care*, 37(3), MS79-MS88.
- Brown, J. (1996). Report on cognitive interviews with Medicaid mothers for the Consumer Assessment of Health Plans Study. DRU-1471-AHCPR, Santa Monica, CA: RAND.
- Cheng, L., & Yang, P. Q. (1996) The "Model Minority" deconstructed. In R. Waldinger & M. Bozorgmehr (eds.), *Ethnic Los Angeles* (pp.305-344). New York, NY: Russell Sage Foundation.
- Coke, E., Rothkopf, E. (1970) Note on a simple algorithm for a computer-produced reading ease score. *Journal of Applied Psychology* 54, 208-210.
- Dale, E. & Chall, J. (1948). A formula for predicting readability: Instructions. *Educational Research Bulletin*, 28, 37-54.
- Dreyer, G. (1984). Readability and responsibility. *Journal of Reading*, 27, 334-339.
- Duffy, T. (1985). Readability formulas: What's the use? In T. Duffy & R. Walker (eds.), *Designing Usable Texts* (pp. 113-143). Orlando, FL: Academic Press, Inc.
- Flesch, R. (1948). A readability yardstick. *Journal of Applied Psychology*, 32, 221-233.

- Fry, E. (1990). A readability formula for short passages. *Journal of Reading*, May, 594-597.
- Fry, E. (1969). The Readability graph validated at primary levels. *The Reading Teacher*, 22, 534-538.
- Fry, E. (1977). Fry's readability graph: Clarifications, validity, and extension to level 17. *Journal of Reading*, 21(3), 242-252.
- Gilliam, B., Pena, S. & Moutain, L. (1980). The Fry graph applied to Spanish readability. *The Reading Teacher*, January, 426-430.
- Gates, A. & MacGinite, W. (1965). *Gates-MacGinite Reading Tests*. New York: Teachers College Press.
- Gunning, R. (1968). The Fog Index after 20 years. *Journal of Business Communication*, 6, 3-13.
- Hibbard, J., Slovic, P. & Jewett J. (1997). Informing consumer decisions in health care: Implications from decision-making research. *The Milbank Quarterly*, 75(3), 395-414.
- Freeman, D. & Freeman, Y. (1997). *Aventuras*. Boston, MA: Houghton Mifflin Company.
- Harris-Kojetin, L., Fowler, F., Brown, J., Schnaier, J. & Sweeny S. (1999). The use of cognitive interviews to develop and evaluate CAHPS®1.0 core survey items. *Medical Care*, 37(3), MS10-MS21.
- Johnston, P. (1983). *Reading Comprehension Assessment: A Cognitive Basis*. Newark, DE: International Reading Association.
- Kayner, G. (1999) Sun Flower. *Cricket*, 26(12), 4-7.

- Kincaid, J., Fishburne, R., Rodgers, R. & Chissom, B. (1975).
Derivation of new readability formulas for Navy enlisted personnel
(Branch Report 8-75). Millington, TN: Chief of Naval Training.
- Kirsch, I., Jungeblut, A., Jenkins, L. & Kolstad, A. (1993). *Adult
Literacy in America*. Princeton, NJ: Educational Testing Service.
- Klare, G. (1974). Assessing readability. *Reading Research Quarterly*, 1,
62-102.
- Lewis, M., Merz, J., Hays, R., & Nicholas, R. (1995). Perceptions of
intoxication and *impairment at arrest among adults convicted of
driving under the influence of alcohol*. *Journal of Drug Issues*, 25,
141-160.
- Leadbetter, C., Hall, S., Swanson, J. & Forrest, K. (1990). Readability
of commercial versus generic health instructions for condoms.
Health Care for Women International, 11, 295-304.
- Lopez, D. (1996). Language and assimilation. In R. Waldinger & M.
Bozorgmehr (eds.), *Ethnic Los Angeles* (pp. 139-163). New York:
Russell Sage Foundation.
- MaCurdy, T. & O'Brien-Strain, M. (1997). *Who Will Be Affected by
Welfare Reform in California?* San Francisco, CA: Public Policy
Institute of California.
- McCall, W. & Crabbs, L. (1961). *Standard Test Lessons in Reading*. New
York: Bureau of Publications, Teachers College, Columbia University.
- McLaughlin, G. (1969). SMOG grading - a new readability formula.
Journal of Reading, 12, 636-646.

Morales, L., Cunningham, W., Brown, J., Lui, H. & Hays, R. (1999). Are Latinos less satisfied with communication by providers? *Journal of General Internal Medicine*, 14, 409-417.

Morrow, G. (1980). How readable are subject consent forms? *JAMA*, 244(1), 56-58.

Romano, R. (1979). *Readability in Cancer Communications: Methods, Examples and Resources for Improving the Readability of Cancer Messages and Materials*. Bethesda, MD: U.S. Department of Health, Education and Welfare, Public Health Service, National Institutes of Health, National Cancer Institute.

Root, J. & Stableford, S. (1999). Easy to read consumer communication: A missing link in medicaid managed care. *Journal of Health Politics, Policy and Law*, 24(1), 1-26.

Rush, R. (1985). Assessing readability: Formulas and alternatives. *Reading Teacher*, 39, 274-283.

Vari-Cartier, P. (1981). Development and validation of a new instrument to assess the readability of Spanish prose. *Modern Language Journal*, 65(Summer), 141-148.

Weech-Maldonado, R., Weidmer, B. O., Morales, L. S. & Hays, R. D. (In Press) Cross-cultural adaptation of survey instruments: The CAHPS® experience. In D. O'Rourke (ed.), *Health Survey Research Methods: Seventh Conference Proceedings*.

Weidmer, B., Brown, J., Garcia, L. (1999). Translating the CAHPS 1.0 Survey Instruments into Spanish. *Medical Care*, 37(3), MS89-MS96.

TABLE 1. Syntactic and Semantic Factor Counts Used in Readability Assessment.

	Syntactic (sentence) Factor	Semantic (word) Factor
Average number of sentences	√	
Average number of words per sentence	√	
Average number of syllables per sentence	√	
Number of characters per sentence	√	
Average number of syllables		√
Average number of 1-syllable words		√
Average number of 2-syllable words		√
Average number of characters per word		√
Average number of syllables per word		√

TABLE 2. Readability Levels of English Language CAHPS- 2.0 Surveys, English Language Newspaper and English Language Children's Book.

	CAHPS® 2.0 Adult Core	CAHPS® 2.0 Child Core	New York Times Article	Cricket Reader (Ages 9 and up)
Fry Readability Graph Score	7 th Grade	7 th Grade	12 th Grade	5 th Grade
Average number of sentences per 100-word sample	5.1	7.9	3.4	9.0
Average number of syllables per 100-word sample	134.0	124.3	153.3	133.0
Flesch Reading Ease Score	71.3	89.6	45.8	81.3
FOG Readability Score	8 th Grade	6 th Grade	12 th Grade	5 th Grade
SMOG Readability Score	7 th Grade	7 th Grade	12 th Grade	7 th Grade
Syntax Indexes				
Average number of words per sentence	19.8	15.2	30.7	11.4
Average number of syllables per sentence	26.5	18.9	46.4	15.1
Average number of characters per sentence	81.2	61.3	141.7	49.9
Semantic Indexes				
Average number of 1-syllable words per 100 words	76.3	83.7	65.3	75.7
Average number of 2- or more syllable words per 100 words	23.7	16.3	34.7	24.3
Average number of characters per word	4.1	4.1	4.7	4.4
Average number of syllables per word	1.3	1.2	1.5	1.3

Note. Fry Readability Graph and Flesch Reading Ease Score based on three 100-word passages taken from the beginning, middle and end of

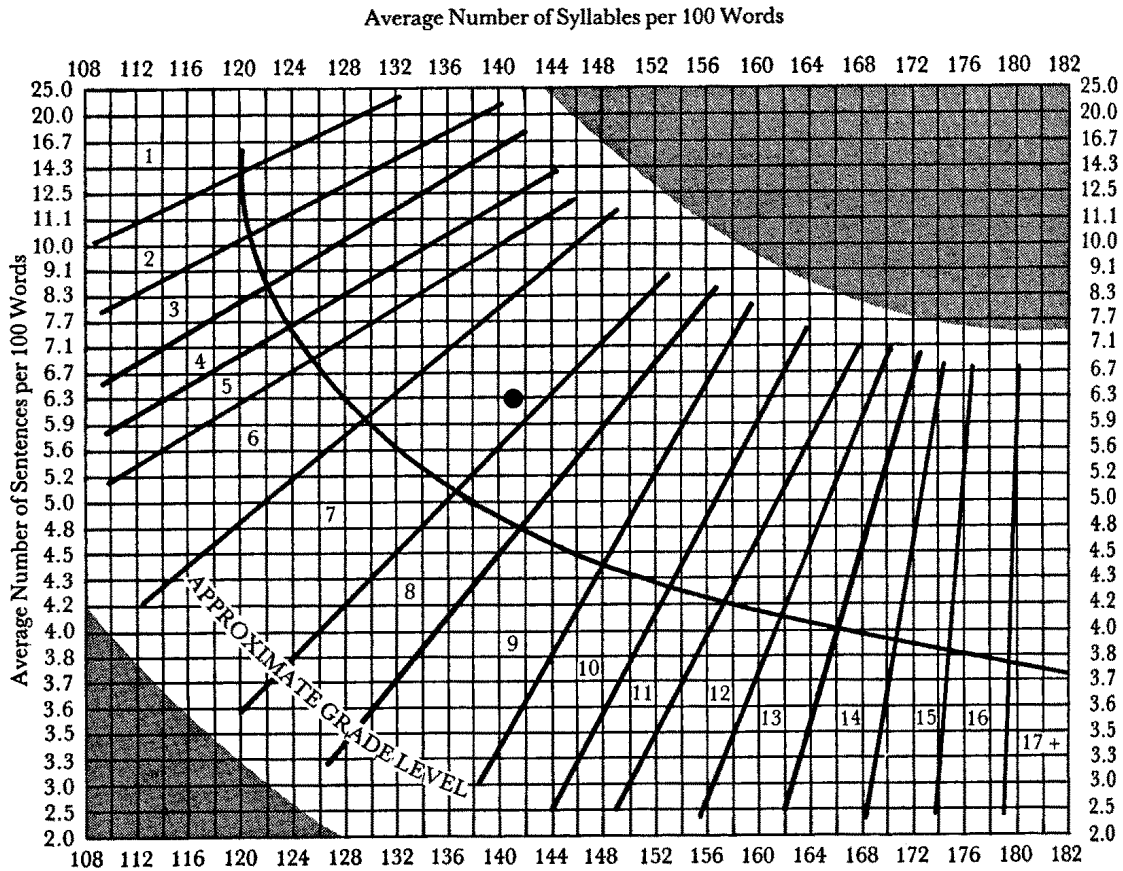
each document. The SMOG score is based on 3 continuous 10-sentence samples taken from the beginning, middle and end of each document.

**TABLE 3. Readability Levels of Spanish Language CAHPS- 2.0 Surveys,
Spanish Language Newspaper and Spanish Language Children's Book.**

	CAHPS® 2.0 Adult Core	CAHPS® 2.0 Child Core	La Opinion	Aventuras
Fry Readability Graph	7 th Grade	7 th Grade	14 th Grade	1 st Grade
Average number of sentences per 100-word samples	6.8	4.4	2.8	16.7
Average number of syllables per 100-word samples	202.7	194.3	235.0	195.7
FRASE Graph	Intermediate	Intermediate	Advanced	Beginning
Syntax Indexes				
Average number of words per sentence	15.6	24.0	38.0	6.0
Average number of syllables per sentence	31.4	46.5	88.2	11.9
Average number of characters per sentence	74.0	110.5	191.6	26.0
Semantic Indexes				
Average number of characters per word	4.8	4.6	5.0	4.3
Average number of syllables per word	2.0	1.9	2.4	2.0

Note. Fry Readability Graph score based on three 100-word passages taken from the beginning, middle and end of each document. The FRASE assessment is based on 5 100-word samples taken from the document.

FIGURE 1
Graph for Estimating Readability—Extended
by Edward Fry, Rutgers University Reading Center, New Brunswick, NJ 08904



DIRECTIONS: Randomly select 3 one hundred word passages from a book or an article. Plot average number of syllables and average number of sentences per 100 words on graph to determine the grade level of the material. Choose more passages per book if great variability is observed and conclude that the book has uneven readability. Few books will fall in gray area but when they do grade level scores are invalid.

Count proper nouns, numerals and initializations as words. Count a syllable for each symbol. For example, "1945" is 1 word and 4 syllables and "IRA" is 1 word and 3 syllables.