



Published in final edited form as:

*Patient Educ Couns.* 2012 October ; 89(1): 69–75. doi:10.1016/j.pec.2012.06.018.

## The Cancer Message Literacy Tests: Psychometric Analyses and Validity Studies

Kathleen M. Mazor, EdD<sup>a</sup>, H. Jane Rogers, PhD<sup>b</sup>, Andrew E. Williams, PhD<sup>c</sup>, Douglas W. Roblin, PhD<sup>d</sup>, Bridget Gaglio, PhD<sup>e</sup>, Terry S. Field, DSc<sup>a</sup>, Sarah M. Greene, MPH<sup>f</sup>, Paul K. J. Han, MD, MPH<sup>g</sup>, and Mary E. Costanza, MD<sup>h</sup>

<sup>a</sup>Meyers Primary Care Institute; University of Massachusetts Medical School; Reliant Medical Group; Fallon Community Health Plan, Worcester, MA, USA

<sup>b</sup>University of Connecticut, Storrs, CT, USA

<sup>c</sup>Kaiser Permanente, Honolulu, HI, USA

<sup>d</sup>Kaiser Permanente, Atlanta, GA, USA

<sup>e</sup>Mid-Atlantic Permanente Research Institute/Kaiser Permanente Mid-Atlantic States, Rockville, MD, USA

<sup>f</sup>Group Health Research Institute, Seattle, WA, USA

<sup>g</sup>Center for Outcomes Research and Evaluation, Portland, ME, USA

<sup>h</sup>University of Massachusetts Medical School, Worcester, MA, USA

### Abstract

**Objectives**—To examine the psychometric properties of two new health literacy tests, and to evaluate score validity.

**Methods**—Adults aged 40 to 71 completed the Cancer Message Literacy Test-Listening (CMLT-Listening), the Cancer Message Literacy Test-Reading (CMLT-Reading), the REALM, the Lipkus numeracy test, a brief knowledge test (developed for this study) and five brief cognitive tests. Participants also self-reported educational achievement, current health, reading ability, ability to understand spoken information, and language spoken at home.

**Results**—Score reliabilities were good (CMLT-Listening: alpha = .84) to adequate (CMLT-Reading: alpha = .75). Scores on both CMLT tests were positively and significantly correlated with scores on the REALM, numeracy, cancer knowledge and the cognitive tests. Mean CMLT scores varied as predicted according to educational level, language spoken at home, self-rated health, self-reported reading, and self-rated ability to comprehend spoken information.

© 2012 Elsevier Ireland Ltd. All rights reserved.

**Corresponding author at:** Kathleen M. Mazor, Meyers Primary Care Institute, 630 Plantation Street, Worcester, MA, USA 01605, Phone: 508.791.7392, Fax: 508.595.2200, Kathy.Mazor@meyersprimary.org, Kathleen.mazor@umassmed.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

#### Conflicts of Interest

Sarah Greene has co-developed free, public domain trainings and tools in the area of health literacy. These tools are freely available online at <http://prism.grouphealthresearch.org>. Her time for developing these tools was funded by a grant from the National Institutes of Health. No additional conflicts of interest were reported by any of the authors.

**Conclusions**—The psychometric findings for both tests are promising. Scores appear to be valid indicators of comprehension of spoken and written health messages about cancer prevention and screening.

**Practice Implications**—The CMLT-Listening will facilitate research into comprehension of spoken health messages, and together with the CMLT-Reading will allow researchers to examine the unique contributions of listening and reading comprehension to health-related decisions and behaviors.

## Keywords

Health literacy; Cancer prevention; Psychometrics

---

## 1. Introduction

Health literacy is increasingly recognized as important,(1–3) and numerous studies have shown links between health literacy, health-related behaviors and health outcomes.(4–14) With few exceptions,(15–20) most research on health literacy has utilized measures that assess print literacy or even more narrowly, word recognition.(21–22) Health literacy is a broader construct, and includes facility with spoken health information.(4, 23–25) This is important because so much health information is transmitted orally.(26) However, without instruments to assess spoken health literacy, there has been very little research on this facet of health literacy.(13) To fill this critical gap, we developed a test to assess comprehension of spoken health messages about cancer prevention and screening, the Cancer Message Literacy Test-Listening (CMLT-Listening). To facilitate future research involving both spoken and written health messages, we developed a companion test assessing comprehension of written messages, the Cancer Message Literacy Test-Reading (CMLT-Reading). Our test development process is summarized in Table 1 and reported in detail in a separate manuscript.(27)

The purpose of the current study was to examine the psychometric properties of the two tests, and to investigate score reliability and validity. We reviewed prior research on health literacy in order to predict associations between CMLT scores and a variety of other measures. Prior research has found lower scores on health literacy measures to be associated with less formal education;(28–31) lower scores on cognitive measures;(32–36) speaking a language other than English at home;(37–38) older age;(28, 33, 37–41) and worse self-rated health.(42–44) We therefore predicted that scores on the two new health literacy measures would show similar associations. We also predicted that the CMLT-Listening and the CMLT-Reading scores would be positively correlated, that each would be correlated with cancer knowledge and numeracy scores, that CMLT-Reading scores would be positively correlated with REALM scores and self-reported reading ability, and that CMLT-Listening scores would be related to self-reported understanding of spoken information.

## 2. Methods

### 2.1 Sites

This study was conducted within the Cancer Research Network (CRN), a consortium of research organizations affiliated with non-profit integrated healthcare delivery systems and the National Cancer Institute (NCI). Four CRN sites participated: Kaiser Permanente Georgia (KPGA), Kaiser Permanente Hawaii (KPHI), Kaiser Permanente Colorado (KPCO) and Fallon Community Health Plan (FCHP) in Massachusetts. The study was approved by the Institutional Review Board at each site.

## 2.2 Participants

Potential participants were identified from health system records by randomly sampling members aged 40–70 who had been enrolled for at least five years, and lived or received care in reasonable proximity to the study session locations. Some participants were aged 71 by the time the study sessions occurred. We targeted this age range because these adults are most likely to face cancer screening decisions, and to be at elevated risk for most cancers compared to younger adults. To optimize sampling across educational levels, at FCHP, KPGA and KPHI sampling was stratified by United States Census-based estimates of educational level defined by the percentage of residents with a high school education or less in the census tract in which participants lived. At KPGA, sampling was further stratified according to the percent of African-American residents, to ensure that African-American and white members were invited in equal numbers within each educational strata. At KPCO, which only recruited Hispanic health plan members, health care system data on race/ethnicity and language preference was used to identify individuals who met the above criteria, self-identified as Latino and had English as their preferred language. A variety of recruitment techniques were used, including mailings, telephone follow-up, and offering sessions at multiple locations. Interested participants were screened to confirm ability to communicate in English, adequate corrected hearing and vision, and the absence of physical or psychological limitations that would preclude participation.

## 2.3 Data collection procedures

Study sessions lasted approximately 2 hours, and were conducted in-person by a trained research assistant. All items (except reading items), were administered orally. Participants provided written informed consent and received \$50 for participation.

## 2.4 Measurements

The CMLT-Listening assesses comprehension of spoken messages related to cancer prevention and screening, and is self-administered via computer. It begins with a brief, computer-narrated introduction, instructions and sample items. The test includes 15 spoken messages presented in video, each with 2 to 4 associated items (48 items total). Items were developed using the sentence verification technique (SVT)(45) and are paraphrases of video message content; participants indicate whether the item meaning is the same as the original message. Details of the test development process are described in an earlier paper.(27) A sample item is provided in Figure 1. Videos cannot be replayed; items can be. Test administration takes approximately 1 hour. No reading is required.

The CMLT-Reading assesses comprehension of written messages on cancer prevention and screening. It is self-administered on paper. It contains 6 messages, each with 3–4 associated items (23 items total). For each item, the participant must indicate whether a statement has the same meaning as the original message. CMLT-Reading items were also developed using the SVT (45) and details of test development are described in a prior paper.(27) A sample item is provided in Figure 2. Administration time is approximately 10 minutes.

Participants also completed the REALM,(21) the Lipkus numeracy test (risk items),(46) a four item test of cancer knowledge developed for this study (appendix), five brief cognitive tests (backward counting, category fluency, word list recall, short delay word list recall and digits backwards),(47) and the five item Perceived Efficacy in Patient Provider Interaction measure (PEPPI).(48) Participants self-reported age, educational attainment, current health, (49) and language spoken at home. Difficulty understanding spoken information was assessed using an item developed specifically for this study (“I have a hard time understanding when people speak quickly”; response options were: strongly disagree, disagree, agree and strongly agree). An item was also developed to assess participants’

perceptions of their reading ability; this item read “I am a good reader”; response options were: strongly disagree, disagree, agree and strongly agree). Participants also indicated whether they had ever been diagnosed with cancer, and whether they were aware of any family member having been diagnosed with cancer.

## 2.5 Analyses

We began by performing dimensionality analyses to determine whether each test was measuring a single construct, as hypothesized. We then computed item, test, and person statistics, and assessed score reliability using coefficient alpha. Dimensionality and reliability are both aspects of score validity,(50–51) but we sought additional validity evidence as well. Specifically, we examined the relationships between scores on the CMLTs and a variety of measures which have been found to be related to scores on other measures of health literacy, and assessed whether those relationships were in the predicted directions.

Both CMLTs included a “Not Sure” response option. Because preliminary psychometric analyses revealed no advantage to including “Not Sure” responses in scoring, all analyses reported here are based only on correct/incorrect scoring.

## 2.6 Dimensionality and Local Independence

Full-information item factor analysis was conducted using TESTFACT.(52) A unidimensional model was initially specified. Then, for each test, two additional analyses were conducted. First, to investigate whether local dependence (and hence multidimensionality) was present due to the structure of the test (i.e., the fact that items were nested within messages) a bifactor model was fitted in TESTFACT, specifying a general factor and specific factors for each set of items based on the same message. Second, to investigate the presence of additional factors unrelated to the structure of the test, an exploratory factor analysis was conducted. For each of these two additional analyses, we evaluated the improvement in fit associated with increasing the number of factors over the unidimensional model. This was assessed by computing the difference in the chi-square fit statistics for the two models as reported by TESTFACT; this difference is a chi-square statistic itself, with degrees of freedom equal to the difference in degrees of freedom for the two models.

## 2.7 Relationships with Other Variables

We examined the relationship between scores on the CMLT tests and categorical variables using ANOVA or t-tests. We examined the relationship between scores on the CMLTs and continuous variables using correlations. Because we planned a large number of statistical tests to examine our predictions about relationships between CMLT scores and related measures, we used a conservative criterion to assess statistical significance ( $p < .001$ ).

## 3. Results

### 3.1 Sample Characteristics

A total of 1074 adult health plan members participated. Sample characteristics are summarized in Table 2.

### 3.2 Dimensionality and Local Independence

For the CMLT-Listening, three items showed negative discrimination estimates, indicating they were not functioning in the same way as the rest of the items on the test; for the CMLT-Reading two items showed negative discrimination estimates. These items were omitted

from all further analyses. For dimensionality analyses, only complete data were used (N = 999).

### 3.3 CMLT-Listening

For the CMLT-Listening, the bifactor model produced 16 factors (1 general and 1 specific for each of the 15 vignettes). The bifactor model yielded a marginal improvement in fit over the unidimensional model ( $\chi^2$  change = 66.2, df = 45, p = .02). Under the bifactor model, the general factor accounted for 21% of the variance, and the 15 specific factors combined accounted for about 9% of the variance, with no specific factor accounting for more than 1% of the variance, suggesting the structure of the test did not create any substantial multidimensionality. Under the simple unidimensional model, the first factor also accounted for 21% of the variance.

To investigate the possibility of a second dimension unrelated to the testlet structure, a two-dimensional exploratory solution was fitted. For the CMLT-Listening, there was a statistically significant improvement in fit with the two-dimensional model ( $\chi^2$  change = 415.9, df = 44, p = .00), with the first and second factors accounting for 22% and 5% of the variance, respectively. However, the second factor appeared to be an artifact related to the difficulty of the items; the correlation between the item loadings on the second factor and the proportion correct value was .73 for the CMLT-Listening. Items that loaded more heavily on the second factor tended to be the easier items, which generally had lower discrimination indices. Given this finding, the lack of any substantive explanation, and the small proportion of variance explained by the factor, a unidimensional model appeared to be a reasonable choice for the data. The relatively small percent of variance accounted for by the primary dimension may be due to the fact that many of the items were very easy with low discriminations; for 26 of the 45 CMLT-Listening items, more than 80% of respondents answered the item correctly, and for 13 of the items, the item-total correlation was less than .2. Thus these items provided little information for model-fitting purposes, with the apparent result that the variance in item scores could not be attributed to the primary dimension.

### 3.4 CMLT-Reading

For the CMLT-Reading, the bifactor model produced a significant improvement in fit ( $\chi^2$  change = 48.5, df = 21, p = .001). The general factor accounted for 30% of the variance, and the 6 specific factors combined accounted for about 8% of the variance, indicating that the improvement in fit provided by the bifactor model was of small practical consequence. Again, the relative weakness of the primary dimension may be due to the lack of variability in the item responses: the CMLT-Reading was very easy for this sample, with over three-quarters of respondents scoring 80% or better.

A two-dimensional exploratory factor model produced a statistically significant improvement in fit over the unidimensional model ( $\chi^2$  change = 155.1, df = 20, p = .00), with the first and second factors accounting for 29% and 6% of the variance, respectively. The second factor was defined largely by four items that had no apparent connection, either psychometrically or substantively; unlike the CMLT-Listening, the second factor did not appear to be strongly related to item difficulty. However, the factor was uninterpretable from a substantive perspective and it explained only a small proportion of variance, so it seemed reasonable to proceed with a unidimensional model.

Finally, local dependence among items based on a common message was assessed by computing the average Q3 statistic for pairs of items in each subset. All averages for both

tests were close to zero, indicating no substantial local dependence among items on either test.(53)

### 3.5 Item and Test Statistics

For the CMLT-Listening, item difficulties (i.e., the proportion of respondents choosing the correct response) ranged from .51 to .97 (mean =.79); for the CMLT-Reading difficulties ranged from .55 to .94 (mean =.84). The mean percent correct total score for the CMLT-Listening (based on 45 items) was 79 (SD=13.6; range 33–100); coefficient alpha was .84; for the CMLT-Reading the total score (based on 21 items) was 84 (SD=14.1; range 24–100); and coefficient alpha was .75. The score distributions were strongly negatively skewed for both tests, indicating that the test was quite easy for most individuals.

### 3.6 Relationships with Other Variables

The correlation between the CMLT-Listening and the CMLT-Reading was .67 ( $p<.001$ ). Other correlations are presented in Table 3. The correlation between the CMLT-Listening and age (in years) was  $-.07$  ( $p=.017$ ); the correlation between age and the CMLT-Reading was .02 ( $p=.45$ ). Correlations between age and scores on the REALM and the numeracy test were not statistically significant. As predicted, less educated participants scored lower on both tests than more educated participants. Relationships with other characteristics were also in the predicted direction (Table 4). Participants who responded “yes” to the question “Have you ever been told by a doctor that you had cancer?” scored slightly but not significantly higher on the CMLT-Listening ( $p=.09$ ) and the CMLT-Reading ( $p=.07$ ). Comparisons of those who reported that a family member had had cancer to those who had not found no evidence of differences on either measure; both  $p$  values were greater than .10.

## 4. Discussion and Conclusion

### 4.1 Discussion

The CMLT-Listening assesses an important but understudied facet of health literacy; comprehension of spoken information related to cancer prevention and screening. The CMLT-Reading provides a companion test which assesses comprehension of print information. Both tests utilize realistic health messages that adults may encounter in their day-to-day lives.

Berkman and colleagues, in a recent systematic review of the literature relating health literacy to health outcomes included an explicit caution that none of the studies they identified examined the relationship between *oral* literacy and outcomes.(13) Other authors have also raised concerns about the limitations of existing instruments.(54,55) This is in spite of the fact that almost all definitions of health literacy, including the widely cited definition used in the landmark IOM report, encompass both oral and print literacy,(4) and recent discussions of health literacy highlight that effective oral communication is key component of health literate care.(56) The CMLT-Listening is intended to fill the critical gap between the widespread recognition of the importance of oral literacy, and meaningful research in this area.

The psychometric findings for these two new measures are promising. Each test is essentially unidimensional. The reliability of the CMLT-Listening scores was .85; the lower but acceptable reliability for the CMLT-Reading scores was not unexpected given the shorter test length.

Scores on the CMLT-Listening and the CMLT-Reading appear to be valid indicators of adults’ comprehension of spoken and written health messages about cancer prevention and

screening. Validity is “the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests.”(50) Establishing validity is a process, not an endpoint. The process begins during test development, and includes careful definition of test content, selection of appropriate stimulus materials and items, and expert review, steps often considered under the label of content validity. We used a careful test development process to create the CMLT-Listening and the CMLT-Reading. The factor analyses results provide further evidence of validity, confirming a single underlying construct for each measure. In addition, the relationships between scores on the CMLTs and other measures suggest that the tests assess important aspects of health literacy. CMLT scores were positively correlated with scores on the REALM, a numeracy test, a brief cancer knowledge test, and five brief cognitive tests. CMLT scores varied as predicted as a function of educational achievement, self-rated health, self-reported reading ability, and self-reported difficulty with spoken communication. Overall, relationships between these two new measures and the variety of established measures and participant characteristics are highly consistent with prior studies of health literacy using other instruments. We found no evidence of a relationship between age and the CMLT-Reading score, the REALM score, or the numeracy score, perhaps due to the restricted age range of our sample (all participants were between the ages of 40 and 71). A recent study of older adults also reported no correlation between health literacy and age.(57)

The correlation between the CMLT-Listening and the CMLT-Reading in this sample was .67, indicating that those who do well on one measure tend to do well on the other. The positive correlation between scores on the listening and reading literacy tests for the adults in this sample may cause some readers to question whether the information provided by the CMLT-Listening is worth the time required to administer it, or whether the CMLT-Reading could be used alone. However, some participants' scores did not covary in this way, suggesting relative strengths or weaknesses in one mode or the other. Additional research is needed to investigate whether such profiles are predictive of communication preferences or health behaviors.

The content domain of both the CMLT-Listening and the CMLT-Reading is cancer prevention and screening. The decision to limit the content was deliberate, based on the premise that people are more or less able to comprehend information in specific domains due to differential experience or interest in certain health conditions. This content focus is intended to facilitate study of cancer prevention and screening specifically. In future studies, we will examine the relationship between CMLT scores and actual cancer screening utilization.

One limitation of the present study is that all participants were adults aged 40 to 71. We chose to limit our sample to participants in this age range due to our focus on cancer prevention and screening, a topic which we anticipated would be most salient to adults in this age range, where cancer screening decisions are most likely to arise. However, one result of this decision is that we are not able to assess the extent to which our findings are generalizable to younger or older adults. A limitation of the CMLT-Listening is that it takes an hour to administer, in part because spoken messages take longer to communicate than print messages. This time requirement may preclude use of this instrument in clinical settings, but researchers studying spoken communication may find it worthwhile to use this test in spite of the time required. In the future, we hope to develop an adaptive and a short form, thereby reducing administration time. In the meantime, the CMLT-Listening provides a standard which shorter tests, screening measures, and subjective-ratings can be compared to. This will facilitate the study of the relationship between print and spoken literacy, and the relative importance of these in healthcare decision making and health behavior.

## 4.2. Conclusions

The CMLT-Listening and the CMLT-Reading are two new measures intended to assess comprehension of spoken and written cancer prevention messages. The CMLT-Listening is the first test that we are aware of to focus specifically and solely on health literacy with respect to spoken information. The psychometric properties of both tests are acceptable, and scores are related to various other measures in predicted directions. The availability of this pair of tests will help to extend our understanding of health literacy, enabling new research into the relative and unique contributions of spoken and print health literacy.

## 4.3 Practice Implications

We hope that the availability of the CMLT-Listening will accelerate research focused on comprehension of spoken health messages. The availability of these two complementary new measures should facilitate research into the unique contributions of listening and reading comprehension to health-related decisions and behaviors.

## Acknowledgments

**Funding/Support and Role of Sponsor:** The study was conducted within the context of a core project of the HMORN Cancer Research Network, *Health Literacy and Cancer Prevention: Do People Understand What They Hear?* funded by the National Cancer Institute (U19 CA079689). Additional funding for participation of a fourth Cancer Research Network site was provided by a pilot grant through the Cancer Research Network Cancer Communication Research Center, also funded by the National Cancer Institute (P20 CA137219). The funding agency did not contribute to the study design, data collection, analysis, or interpretation; or the decision to submit the manuscript for publication.

**Acknowledgement:** The research team would like to extend a sincere thank you to Melissa Finucane, PhD, for her efforts at the start of the project. The authors would like to thank Joann Wagner, Erica Cove, and Laura Saccoccio for their help with project design and data collection. The authors wish to thank Mallory Thomas for her assistance in preparing and formatting the manuscript. The authors would also like to acknowledge *JAMA*® for the generous contribution of video content related to HPV infection for the CMLT-Listening.

## References

1. U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion. National action plan to improve health literacy. Washington DC: Department of Health and Human Services; 2010.
2. Benjamin R. Health literacy improvement as a national priority. *J Health Commun.* 2010; 15 1-
3. Carmona RH. Health literacy: a national priority. *J of Gen Int Med.* 2006; 21:803.
4. Nielsen-Bohlman, L.; Panzer, AM.; Kindig, D. Health literacy: a prescription to end confusion. Washington, DC: The National Academies Press; 2004.
5. Rudd, RE.; Moeckens, BA.; Colton, TC. Health and literacy. A review of medical and public health literature. In: Comings, JGB.; Smith, C., editors. Annual review of adult learning and literacy. New York: Jossey-Bass; 2000.
6. Berkman ND, Dewalt DA, Pignone MP, Sheridan SL, Lohr KN, Lux L, Sutton SF, Swinson T, Bonito AJ. Literacy and health outcomes. *Evid Rep Technol Assess (Summ).* 2004;1–8. [PubMed: 15819598]
7. Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs, American Medical Association. Health literacy: report of the Council on Scientific Affairs. *J Amer Med Assoc.* 1999; 281:552–557.
8. Baker DW, Parker RM, Williams MV, Clark WS. Health literacy and the risk of hospital admission. *J Gen Intern Med.* 1998; 13:791–798. [PubMed: 9844076]
9. Sudore RL, Yaffe K, Satterfield S, Harris TB, Mehta KM, Simonsick EM, Newman AB, Rosano C, Rooks R, Rubin Sm, Ayonayon HN, Schillinger D. Limited literacy and mortality in the elderly: the health, aging, and body composition study. *J Gen Intern Med.* 2006; 21:806–812. [PubMed: 16881938]



10. von Wagner C, Steptoe A, Wolf MS, Wardle J. Health literacy and health actions: a review and a framework from health psychology. *Health Educ Behav*. 2009; 36:860–877. [PubMed: 18728119]
11. Dewalt DA, Berkman ND, Sheridan S, Lohr KN, Pignone MP. Literacy and health outcomes: a systematic review of the literature. *J Gen Intern Med*. 2004; 19:1228–1239. [PubMed: 15610334]
12. Williams MV, Baker DW, Parker RM, Nurss JR. Relationship of functional health literacy to patients' knowledge of their chronic disease. A study of patients with hypertension and diabetes. *Arch Intern Med*. 1998; 158:166–172. [PubMed: 9448555]
13. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. *Ann Intern Med*. 2011; 155:97–107. [PubMed: 21768583]
14. Berkman, ND.; Sheridan, SL.; Donahue, KE.; Halpern, DJ.; Viera, A.; Crotty, K.; Holland, A.; Brasure, M.; Lohr, KN.; Harden, E.; Tant, E.; Wallace, I.; Viswanathan, M. Health literacy interventions and outcomes: an updated systematic review. Rockville, MD: Agency for Healthcare Research and Quality; 2011.
15. Koch-Weser S, Rudd RE, Dejong W. Quantifying word use to study health literacy in doctor-patient communication. *J Health Commun*. 2010; 15:590–602. [PubMed: 20812122]
16. Martin LT, Schonlau M, Haas A, Derose KP, Rudd R, Loucks EB, Rosenfield L, Buka SL. Literacy skills and calculated 10-year risk of coronary heart disease. *J Gen Intern Med*. 2011; 26:45–50. [PubMed: 20809155]
17. Mazor KM, Calvi J, Cowan R, Costanza ME, Han PK, Greene SM, Saccoccio L, Cove E, Roblin D, Williams A. Media messages about cancer: what do people understand? *J Health Commun*. 2010; 15:126–145. [PubMed: 20845199]
18. Rosenfeld L, Rudd R, Emmons KM, Acevedo-Garcia D, Martin L, Buka S. Beyond reading alone: the relationship between aural literacy and asthma management. *Patient Educ Couns*. 2011; 82:110–116. [PubMed: 20399060]
19. Roter DL, Erby L, Larson S, Ellington L. Oral literacy demand of prenatal genetic counseling dialogue: predictors of learning. *Patient Educ Couns*. 2009; 75:392–397. [PubMed: 19250792]
20. McCormack L, Bann C, Squiers L, Berkman ND, Scquire C, Schillinger D, Ohene-Frempong J, Hibbard J. Measuring health literacy: a pilot study of a new skills-based instrument. *J Health Commun*. 2010; 15:51–71. [PubMed: 20845193]
21. Davis TC, Long SW, Jackson RH, Mayeaux EJ, George RB, Murphy PW, Crouch MA. Rapid estimate of adult literacy in medicine: a shortened screening instrument. *Fam Med*. 1993; 25:391–395. [PubMed: 8349060]
22. Parker RM, Baker DW, Williams MV, Nurss JR. The test of functional health literacy in adults: a new instrument for measuring patients' literacy skills. *J Gen Intern Med*. 1995; 10:537–541. [PubMed: 8576769]
23. Berkman ND, Davis TC, McCormack L. Health literacy: what is it? *J Health Commun*. 2010; 15:9–19. [PubMed: 20845189]
24. Williams M, Davis T, Parker R, Weiss B. The role of health literacy in patient-physician communication. *Fam Med*. 2002; 34:383–389. [PubMed: 12038721]
25. Rudd RE. Improving Americans' health literacy. *N Engl J Med*. 2010; 363:2283–2285. [PubMed: 21142532]
26. Gallup Organization. Gallup poll. Storrs, CT: Roper Center for Public Opinion Research; 2002.
27. Mazor KM, Roblin DW, Williams AE, Greene SM, Gagilo B, Field TS, Costanza ME, Han PK, Saccoccio L, Calvi J, Cove E, Cowan R. Health literacy and cancer prevention: two new instruments to assess comprehension. *Patient Educ Couns*. 2012; 88:54–60. [PubMed: 22244323]
28. Ginde AA, Weiner SG, Pallin DJ, Camargo CA Jr. Multicenter study of limited health literacy in emergency department patients. *Acad Emerg Med*. 2008; 15:577–580. [PubMed: 18616448]
29. Morrow D, Clark D, Tu W, Wu J, Weiner M, Steinley D, Murray MD. Correlates of health literacy in patients with chronic heart failure. *Gerontologist*. 2006; 46:669–676. [PubMed: 17050758]
30. Wister AV, Malloy-Weir LJ, Rootman I, Desjardins R. Lifelong educational practices and resources in enabling health literacy among older adults. *J Aging Health*. 2010; 22:827–854. [PubMed: 20595098]

31. Martin LT, Ruder T, Escarce JJ, Ghosh-Dastidar B, Sherman D, Elliot M, Bird CE, Fremont A, Gasper C, Culbert A, Lurie N. Developing predictive models of health literacy. *J Gen Intern Med.* 2009; 24:1211–1216. [PubMed: 19760299]
32. Baker DW, Gazmararian JA, Sudano J, Patterson M, Parker RM, Williams MV. Health literacy and performance on the Mini-Mental State Examination. *Aging Ment Health.* 2002; 6:22–29. [PubMed: 11827619]
33. Gazmararian JA, Baker DW, Williams MV, Parker RM, Scott TL, Green DC, Fehrenbach SN, Ren J, Koplan JP. Health literacy among Medicare enrollees in a managed care organization. *J Amer Med Assoc.* 1999; 281:545–551.
34. Federman AD, Sano M, Wolf MS, Siu AL, Halm EA. Health literacy and cognitive performance in older adults. *J Am Geriatr Soc.* 2009; 57:1475–1480. [PubMed: 19515101]
35. Wolf MS, Gazmararian JA, Baker DW. Health literacy and functional health status among older adults. *Arch Intern Med.* 2005; 165:1946–1952. [PubMed: 16186463]
36. Levinthal BR, Morrow DG, Tu W, Wu J, Murray MD. Cognition and health literacy in patients with hypertension. *J Gen Intern Med.* 2008; 23:1172–1176. [PubMed: 18459011]
37. Olives T, Patel R, Patel S, Hottinger J, Miner JR. Health literacy of adults presenting to an urban ED. *Am J Emerg Med.* 2011; 29:875–882. [PubMed: 20627212]
38. Williams MV, Parker RM, Baker DW, Parikh NS, Pitkin K, Coates WC, Nurss JR. Inadequate functional health literacy among patients at two public hospitals. *J Amer Med Assoc.* 1995; 274:1677–1682.
39. Baker DW, Wolf MS, Feinglass J, Thompson JA, Gazmararian JA, Huang J. Health literacy and mortality among elderly persons. *Arch Intern Med.* 2007; 167:1503–1509. [PubMed: 17646604]
40. Baker DW, Gazmararian JA, Sudano J, Patterson M. The association between age and health literacy among elderly persons. *J Gerontol B Psychol Sci Soc Sci.* 2000; 55:S368–S374. [PubMed: 11078114]
41. Montalto NJ, Spiegler GE. Functional health literacy in adults in a rural community health center. *W V Med J.* 2001; 97:111–114. [PubMed: 11392190]
42. Baker DW, Parker RM, Williams MV, Clark WS, Nurss J. The relationship of patient reading ability to self-reported health and use of health services. *Am J Public Health.* 1997; 87:1027–1030. [PubMed: 9224190]
43. Bennett IM, Chen J, Soroui JS, White S. The contribution of health literacy to disparities in self-rated health status and preventive health behaviors in older adults. *Ann Fam Med.* 2009; 7:204–211. [PubMed: 19433837]
44. Baker DW, Gazmararian JA, Williams MV, Scott T, Parker RM, Green D, Ren J, Peel J. Functional health literacy and the risk of hospital admission among Medicare managed care enrollees. *Am J Public Health.* 2002; 92:1278–1283. [PubMed: 12144984]
45. Royer J. Developing reading and listening comprehension tests based on the Sentence Verification Technique. *J Adol Adult Literacy.* 2001; 45:30–41.
46. Lipkus IM, Samsa G, Rimer BK. General performance on a numeracy scale among highly educated samples. *Med Decis Making.* 2001; 21:37–44. [PubMed: 11206945]
47. Tun PA, Lachman ME. Telephone assessment of cognitive function in adulthood: the Brief Test of Adult Cognition by Telephone. *Age Ageing.* 2006; 35:629–632. [PubMed: 16943264]
48. Maly RC, Frank JC, Marshall GN, DiMatteo MR, Reuben DB. Perceived efficacy in patient-physician interactions (PEPPI): validation of an instrument in older persons. *J Am Geriatr Soc.* 1998; 46:889–894. [PubMed: 9670878]
49. National Cancer Institute. [accessed 2007 Oct 16] Health Information National Trends Survey (HINTS). 2005. Available from: <http://hints.cancer.gov/questions/section.jsp?section=Health+Status>.
50. The American Educational Research Association, The American Psychological Association, & The National Council on Measurement in Education. Standards for educational and psychological testing. Washington, D.C: American Education Research Association; 1999.
51. Downing SM. Validity: on meaningful interpretation of assessment data. *Med Educ.* 2003; 37:830–837. [PubMed: 14506816]

52. duToit, M. IRT from SSI: BILOG-MG, MULTILOG, PARSCALE, TESTFACT. Lincolnwood, IL: Scientific Software International; 2003.
53. Yen W. Scaling performance assessments: strategies for managing local item independence. *J Educ Meas.* 1993; 30:187–213.
54. Pleasant A, McKinney J, Rikard RV. Health literacy measurement: a proposed research agenda. *J Health Commun.* 2011; 16:11–21. [PubMed: 21951240]
55. Jordan JE, Osborne RH, Buchbinder R. Critical appraisal of health literacy indices revealed variable underlying constructs, narrow content and psychometric weaknesses. *J Clin Epidemiol.* 2011; 64:366–379. [PubMed: 20638235]
56. Koh HK, Berwick DM, Clancy CM, Baur C, Brach C, Harris LM, Zerhusen EG. New federal policy initiatives to boost health literacy can help the nation move beyond the cycle of costly 'crisis care'. *Health Aff (Millwood).* 2012; 31:434–443. [PubMed: 22262723]
57. McCarthy DM, Waite KR, Curtis LM, Engel KG, Baker DW, Wolf MS. What did the doctor say? Health literacy and recall of medical instructions. *Med Care.* 2012; 50:277–282. [PubMed: 22411440]

## Appendix. Cancer Knowledge Items

1. If your doctor tells you that you are “at risk” for getting colon cancer, it means that you will probably get colon cancer.
2. If your doctor tells you a cancer screening test has a high false positive rate, it means that the test often MISSES the early signs of cancer.
3. If your doctor tells you that it’s important to find cancer at an early stage, it means that its best to find it while you are still relatively young, and in the early stages of your life.
4. The reason that the doctor will remove any polyps that are found during a colonoscopy is so that they don’t break loose and run rampant throughout the body.

**Response Options:** True; False; Not Sure

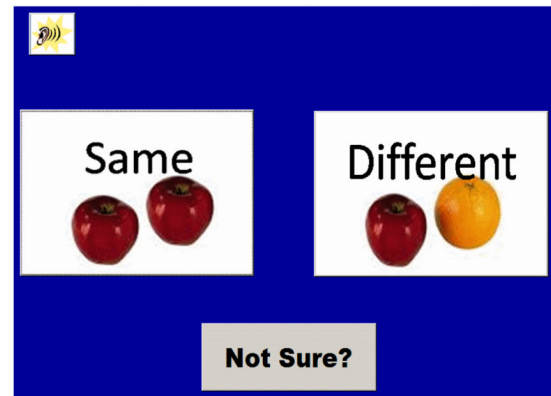
The correct response for all items is “false”.

### Clinical Message



*Physician:* “Well, because your brother had colon cancer, I would recommend the most thorough screening available, which is a colonoscopy.”

### Test Item



“A family history of cancer should not influence what screening procedure you choose.”

“Same meaning or different?”

**Figure 1.**  
CMLT-Listening Test Sample Item

Clinical Message

Test Item

**How can having mammograms help you stay healthy?**  
**From A Patient Education Booklet**

***What is the chance of getting breast cancer?***  
1 in 8 is the *average* chance of a woman getting breast cancer during her lifetime. Your own chance of ever getting breast cancer might be a little lower than 1 in 8. Or it could be a lot higher.

On average, 8 percent of women get breast cancer during their lives.....

<b>SAME</b>	<b>DIFFERENT</b>	<b>NOT SURE</b>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Figure 2.**  
CMLT-Reading Sample Item

\$watermark-text

\$watermark-text

\$watermark-text

**Table 1**

## Overview of the Test Development Process

Test Development Step	Result
Specify Purpose of Measurement	To assess two critical components of health literacy: comprehension of spoken and written health messages.
Specify Content Parameters	Cancer prevention and screening; common cancers and cancers for which screening was available, or for which behavior may influence risk. Realistic and typical of the cancer prevention and screening messages that might be encountered in day-to-day life, either in the media or in clinical settings.
Identify Candidate Messages	Publicly available messages identified via the internet, including posted television and radio clips, health websites, magazines, and newspapers. Clinical materials identified (print) or created (simulated clinician-patient encounters).
Select Messages	Candidate messages reviewed; inaccurate or potentially misleading messages excluded. If multiple messages available, selections made to provide variety of content and type.
Write / Edit Items	Items written by professional writers using the Sentence Verification Technique (45); reviewed and edited by team.
Select Items	Initial item selection sought to include potentially difficult concepts (e.g., risk); and to achieve balance and variety of content.
Pre-test	Pre-testing completed with 7 adults to identify programming errors and to pre-test test administration processes.
Pilot test	Pilot testing conducted with 89 adults from three sites; team examined corrected item-total correlations, coefficient alphas, and differences in the proportion of correct responses across quartiles.
Revise Items and Finalize Item Selection	Poorly performing items edited or replaced; item selection finalized.

**Table 2**

## Sample Characteristics

Characteristic		N	%
Study Site	Colorado	162	15
	Georgia	299	28
	Hawaii	303	28
	Massachusetts	310	29
Gender	Male	439	41
	Female	634	59
Race/Ethnicity	American Indian/Alaskan Native	5	<1
	Black/African American	146	14
	Native Hawaiian/Other Pacific Islander	11	1
	Asian	118	11
	White/Caucasian	519	48
	Hispanic (No Other Race Identified)	163	15
	Hispanic and Other	33	3
	Multiple Races Unknown/Not Reported	55 24	5 2
Language Spoken at Home	English	969	90
	English and Other	76	7
	Other	20	2
Education	Less than High School	32	3
	High School Graduate (includes technical training)	250	23
	Some College or Associates Degree	292	27
	Bachelors Degree	260	24
	Graduate Degree	231	22
Age	40–49	192	18
	50–59	396	37
	60+	475	44
Marital Status	Married	697	65
Work Status	Working for pay	622	58
	Retired	304	28
	Other	139	13
Total Study Participants		1074	100

**Table 3**

Correlations between CMLT Scores and Scores on Related Measures

	<b>CMLT-Listening</b>	<b>CMLT-Reading</b>
REALM	.38	.46
Numeracy	.54	.52
Cancer Knowledge	.55	.43
Word List Recall	.29	.30
Short Delay Word List Recall	.29	.25
Digits Backwards	.31	.32
Category Fluency	.36	.34
Backwards Counting	.37	.32

All correlations in this table are statistically significant;  $p < .001$



**Table 4**

## Mean CMLT Scores by Level of Selected Characteristics

	CMLT-Listening mean score	CMLT-Reading mean score
<b>Educational Attainment</b>	63	68
Less than high school	72	79
High school or trade school	77	82
Some college or Associate Degree	83	87
Four year college degree	86	92
Post-graduate degree		
<b>Language spoken at home</b>	80	85
English only	72	82
Some other language		
<b>Self-rated health</b>	80	87
Very good or Excellent	77	83
Poor, Fair or Good		
<b>"I am a good reader"</b>	80	86
Agree or Strongly Agree	73	78
Disagree or Strongly Disagree		
<b>"I have a hard time understanding when people speak quickly"</b>	81	87
Disagree or Strongly Disagree	76	83
Agree or Strongly Agree		

F statistics for both ANOVAs were statistically significant  $p < .001$ ; all t statistics were statistically significant  $p < .001$  except the t statistic for the CMLT-Reading and Language spoken at home which was  $p < .01$ .