ORIGINAL ARTICLES

The Test of Functional Health Literacy in Adults:

A New Instrument for Measuring Patients' Literacy Skills

Ruth M. Parker, MD, David W. Baker, MD, MPH, Mark V. Williams, MD, Joanne R. Nurss, PhD

OBJECTIVE: **To develop a valid, reliable instrument to measure the functional health literacy of patients.**

DESIGN: **The Test of Functional Health Literacy in Adults (TOFHLA) was developed using actual hospital materials. The TOFHLA consists of a 50-item reading comprehension and 17-item numerical ability test, taking up to 22 minutes to administer. The TOFHLA, the Wide Range Achievement Test–Revised (WRAT-R), and the Rapid Estimate of Adult Literacy in Medicine (REALM) were administered for comparison. A Spanish version was also developed (TOFHLA-S).**

SETTING: **Outpatient settings in two public teaching hospitals.**

PATIENTS: **256 English- and 249 Spanish-speaking patients were approached. 78% of the English- and 82% of the Spanish-speaking patients gave informed consent, completed a demographic survey, and took the TOFHLA or TOFHLA-S.**

RESULTS: **The TOFHLA showed good correlation with the WRAT-R and the REALM (correlation coefficients 0.74 and 0.84, respectively). Only 52% of the English speakers completed more than 80% of the questions correctly. 15% of the patients could not read and interpret a prescription bottle with instructions to take one pill by mouth four times daily, 37% did not understand instructions to take a medication on an empty stomach, and 48% could not determine whether they were eligible for free care.**

CONCLUSIONS: **The TOFHLA is a valid, reliable indicator of patient ability to read health-related materials. Data suggest that a high proportion of patients cannot perform basic reading tasks. Additional work is needed to determine the prevalence of functional health illiteracy and its effect on the health care experience.**

KEY WORDS: **functional health literacy; literacy and health; health literacy measurement.**
J GEN INTERN MED 1995;10:537–541.

Illiteracy is a well-recognized national crisis. Results from the National Adult Literacy Survey, which became available September 1993, provide the most detailed portrait ever available of the condition of literacy in our country. The survey of 13,600 individuals found that 22% of adult Americans, some 40 to 44 million people, perform at the lowest skill level. One of four of these people reported physical, mental, or health conditions that keep them from participating fully in work, school, or housework. One fourth of those in the lowest reading level are immigrants whose native language is not English.[1] The prevalence of low literacy is supported by data from the U.S. Census, which defines illiteracy as those having an eighth-grade education or less. Using that definition of illiteracy, 27 million Americans are illiterate; another 45 million are only marginally literate, meaning up to one of every three adult Americans is functionally illiterate.[2, 3] Functional literacy is the ability to use reading, writing, and computational skills at a level adequate to meet the needs of everyday life situations.

Basic skills in reading, writing, and "numeracy" are especially important in the health care setting, where patient participation in planning and implementing therapeutic regimens is critical for success. Patients need to be able to understand oral and written information about their medical conditions, follow written and numerical directions regarding their therapeutic regimens and diagnostic tests, ask pertinent questions of medical personnel, report prior conditions and treatment, and solve problems that arise during the course of their care. Adequate functional health literacy means being able to apply literacy skills to health-related materials such as prescriptions, appointment cards, medicine labels, and directions for home health care.

Little has been done to explore the impact of illit-

Received from the Emory University School of Medicine (RMP, DWB, MVW), Atlanta, Georgia; the Harbor–UCLA Medical Center (DWB), Los Angeles, California; and the Georgia State University (JRN), Atlanta, Georgia.

Presented at the annual meeting of the Society of General Internal Medicine, Washington, DC, April 27–29, 1994, and the annual meeting of the Robert Wood Johnson Clinical Scholars, November 10, 1994, Fort Lauderdale, Florida.

Supported by a grant from The Robert Wood Johnson Foundation.

Address correspondence to Dr. Parker: Department of Medicine, Emory University School of Medicine, 69 Butler Street, SE, Atlanta, GA 30303. Reprints are not available.

eracy on health care. Published studies indicate that the last grade completed in school is not a good indicator of reading ability[4] and that a significant number of patients have difficulty reading discharge instructions.[5–7] One study has shown illiteracy and poor health status to be independently associated.[8] Studies of functional health literacy have been limited by the lack of an appropriate testing instrument. Literacy assessment tools such as the Wide Range Achievement Test–Revised (WRAT-R)[9] can be used to assign a grade level, but interpreting results is problematic because grade level does not necessarily give an estimate of functional health literacy. The Rapid Estimate of Adult Literacy in Medicine (REALM), which uses only health-related words, has been used to identify a high proportion of indigent outpatients who had poor reading ability.[10] The WRAT-R is not available in Spanish and the REALM is not valid in Spanish.[11] Neither of these instruments tests the ability to read and understand numbers, referred to by literacy experts as *numeracy*, or *quantitative literacy*. Numeracy skills may be the most important element for functional health literacy. To better understand functional health literacy, we developed the Test of Functional Health Literacy in Adults (TOFHLA). The TOFHLA tests a patient's ability to read passages (TOFHLA: Reading Comprehension) and phrases containing numbers (TOFHLA: Numeracy) using real materials from the health care setting. This paper describes how the TOFHLA was developed and the results of initial tests.*

METHODS

For test development, a literacy expert reviewed more than 30 examples of commonly used hospital texts, including patient education materials, instructions for diagnostic tests, prescription bottle labels and instructions, and patient registration forms. The TOFHLA was developed from a sample of these items that were believed to be widely used and of varying difficulties. The test consists of two parts: Reading Comprehension and Numeracy. The Reading Comprehension section is a 50-item test using the modified Cloze procedure[12]; that is, every fifth to seventh word in a passage is omitted. The reader selects from four possible choices, one of which is correct and three of which are similar but grammatically or contextually incorrect. Passages were selected from instructions for preparation for an upper gastrointestinal series, the patient rights and responsibilities section of a Medicaid application form, and a standard hospital informed consent form. The readability levels of the passages on the Gunning Fog index[13] are grades 4.3, 10.4, and 19.5, respectively.

The Numeracy section is a 17-item test using actual hospital forms and labeled prescription vials. It tests a patient's ability to comprehend directions for taking medicines, monitoring blood glucose, keeping clinic appointments, and obtaining financial assistance. Patients are presented with cue cards or labeled prescription bottles and asked to respond to oral questions regarding information about the cards or bottles. The overall readability level of the numeracy prompts on the Gunning Fog index is grade 9.4. The numeracy score is multiplied by 2.941 to create a score from 0 to 50, the same range as that for the reading comprehension scores. The sum of the reading comprehension and the weighted numeracy scores yields the TOFHLA score, which ranges from 0 to 100 and has equal contributions from the two sections.

Item difficulties (p-values and biserial correlations) for each TOFHLA response were calculated based on the responses from the test population. Items were selected to obtain a median difficulty of 72% for reading comprehension and 64% for numeracy. One distractor (an option that is an incorrect answer) was changed for three reading comprehension items and six numeracy items were dropped for the final edition. One numeracy item asked the respondent to choose which of four spoons represented a tablespoon, needed for proper dosing of antacids. Eighty-eight percent of the patients could not correctly answer this question, which was then omitted in the final version of TOFHLA.

To develop a Spanish version of TOFHLA, or TOFHLA-S, the reading comprehension passages and numeracy questions were translated into Spanish and backtranslated into English. Discrepancies were corrected using the consensus of several bilingual staff members and a Spanish literacy expert. The Cloze procedure was then performed on each reading passage to achieve difficulty comparable to that of the English passage instead of using the same word deletions and response options as the English version had.

For test development, pilot studies were conducted at Grady Memorial Hospital and Harbor–UCLA Medical Center in early 1993. Grady Memorial Hospital is an approximately 1,000-bed public hospital in Atlanta, Georgia. The vast majority of patients are African-American indigent residents of Dekalb and Fulton Counties. The nonappointment acute care clinics are the site of more than 320,000 patient visits yearly. A convenience sample of 256 patients presenting for acute care to the medical walk-in clinic and emergency care center were asked by two trained research assistants to participate in the test development study. Eleven percent were excluded due to preestablished criteria, which included age less than 18 years, smell of alcohol on the breath, unintelligible speech, English as a second language, overt psychiatric illness, police custody, illness so severe

**The TOFHLA is copyrighted and is available upon request from: Joanne R. Nurss, PhD, Director, Center for the Study of Adult Literacy, Georgia State University, 1 University Plaza, Atlanta, GA, 30303-3083.*

as to preclude participation, and refusal to participate. The patients' visual acuity was screened using the Rosenbaum hand-held vision chart. Using a criterion of 20/50, 11% of the population failed the screening and were excluded from further testing.

A large-print version has now been developed and is available for future research. Of the patients initially selected, 78% (200/256) gave informed consent and completed demographic inventories, the TOFHLA, the REALM, and the reading subtest of the WRAT-R. The REALM is a list of 66 medically related words, which patients are asked to read aloud to check for correct pronunciation; scores are categorized into one of four literacy levels judged equivalent to grades 0 through 12. The reading subtest of the WRAT-R is a brief screening test in which the participant reads a list of 46 general words and receives a grade-equivalent score. These two measures were given to obtain an estimate of concurrent validity. Three measures of material deprivation were used in the demographic survey to estimate economic status: type of housing, car ownership, and whether the patient was currently receiving any form of public assistance.

Harbor–UCLA Medical Center is a 500-bed public hospital in Torrance, California, that is owned and operated by Los Angeles County. The hospital serves an extremely diverse patient population; approximately 40% are Hispanic, 30% are African-American, 25% are Caucasian, and the remaining 5% are of other ethnic minorities. A convenience sample of native Spanish-speaking patients from the ambulatory care clinics were asked by two trained bilingual research assistants to participate. Similar exclusion criteria were used. Of the 249 patients asked to participate, 203 (82%) gave informed consent and completed the demographic survey and the TOFHLA-S.

Data Analysis

Data were analyzed using SPSS.[14] Reliability was calculated by both split-half and internal consistency measures, using equal-length Spearman-Brown and Cronbach's alpha formulas, respectively. Content validity was enhanced by using actual hospital medical texts for both the Reading Comprehension and the Numeracy subtests. Concurrent validity was tested by determining Spearman's rank correlation between the TOFHLA, the WRAT-R, and the REALM.

RESULTS

The mean ages of the English- and the Spanish-speaking patients were 40 and 42 years, respectively (Table 1). Sixty-eight percent of the patients at Harbor–UCLA were women, which is consistent with the pattern of use among Hispanics at that site. Almost all the English-speaking patients were African-American, which is consistent with the demographics of the overall patient population served by Grady Hospital. Forty-one percent of the English-speaking and 76% of the Spanish-speaking patients had less than a high school education. The measures of economic status showed that only a minority of the patients owned a car, half did not have their own homes or apartments, and a third were receiving public assistance. The English-speaking patients were less likely to own a car and more likely to be receiving some form of public assistance.

Table 1
Patient Characteristics

	English (n = 200)	Spanish (n = 203)
Age—mean	40 years	42 years
Gender—female	51%	68%
Race/ethnicity		
African-American	91%	0%
Caucasian	7%	0%
Hispanic	1%	99%
Education*		
<6th grade	8%	63%
7th to 11th grade	33%	13%
High school graduate/GED†	40%	15%
Some college	21%	9%
Socioeconomic indicators		
Own/rent own residence	56%	53%
Own a car‡	26%	35%
Receive public assistance§	44%	20%

*$p < 0.001$ by chi-square test.
†GED = graduate equivalency diploma.
‡$p = 0.05$ by chi-square test.
§$p < 0.01$ by chi-square test.

Reliability was calculated using all 67 items for both the total TOFHLA and the total TOFHLA-S tests. Both the Cronbach's alpha measure of internal consistency and the Spearman-Brown equal-length coefficient, an estimate of test–retest reliability, were excellent (Table 2). Correlations of the TOFHLA with the REALM and the WRAT-R were 0.84 and 0.74, respectively ($p < 0.001$ by Spearman's rank correlation). The REALM is not valid in Spanish and the WRAT-R is not available in Spanish, so similar correlations with TOFHLA-S were not possible. Intercorrelations among the Reading Comprehension and the Numeracy subtests were $r = 0.79$ and 0.70 for the English and Spanish versions of the TOFHLA, respectively. These correlations demonstrate the internal consistency of the TOFHLA as well as the unique contributions of the two subtests.

The mean TOFHLA score was 68.6 for the English- and 59.0 for the Spanish-speaking patients. Half of the English-speaking and less than a third of the Spanish-speaking patients completed 80% or more of the items correctly (Table 3). Less than 50% of the questions were correctly completed by 27% of the English-speaking and

37% of the Spanish-speaking patients. The mean Numeracy subtest score was 35.0 for the English-speaking and 32.4 for the Spanish-speaking patients. Summary statistics for individual questions showed that the patients struggled with even basic tasks. For example, 15% of the English-speaking patients could not read and interpret a prescription bottle with instructions to take one pill by mouth four times per day, and 37% did not understand instructions to take a medication on an empty stomach. For the Reading Comprehension subtest, the mean score was 33.6 for the English-speaking and 26.6 for the Spanish-speaking patients. Forty-eight percent of the English-speaking and 53% of the Spanish-speaking patients could not determine whether they were eligible for free care.

DISCUSSION

This study reports the development of the TOFHLA, the first available tool for measuring functional health literacy. The TOFHLA measures patients' ability to perform health-related tasks that require reading and computational skills. Literacy experts have shown that functional literacy varies by context and setting. In other words, functional literacy is situation-specific: someone's reading skills may be perfectly adequate in one setting and marginal or inadequate in another. Previously published studies about literacy in the health care setting have focused on patient ability to pronounce a list of words correctly.[7-10] These studies document that many commonly used educational materials and informed consents are often written at levels above many patients' abilities.[4,15] The appropriate tool for measuring functional health literacy should use real and relevant medical texts from our hospitals or clinics. In addition to measuring prose-reading skills, it should measure numeracy skills, which are an essential component of functional health literacy. The TOFHLA is unique in measuring functional health literacy in adults. Having the appropriate tool available to measure functional health literacy will allow future studies to more closely identify the relationship of literacy to the health care experience.

Table 2
Reliability of the Test of Functional Health Literacy in Adults (TOFHLA) and Validity of the TOFHLA when Compared with the Rapid Estimate of Adult Literacy in Medicine (REALM) and the Wide Range Achievement Test–Revised (WRAT-R)

	English (n = 200)	Spanish (n = 203)
Reliability		
Spearman-Brown	0.92	0.84
Cronbach's alpha	0.98	0.98
Validity (r)*		
REALM	0.84†	—
WRAT-R	0.74†	—

*The REALM is not valid in Spanish and the WRAT-R is not available in Spanish.
†$p < 0.001$ by Spearman's rank correlation.

Table 3
Proportion of Items Correctly Completed during Testing of the Test of Functional Health Literacy in Adults (TOFHLA)

	English (%)	Spanish (%)
≥80%	52	31
50–79%	21	32
<50%	27	37

Results from this development study show that the TOFHLA is a reliable tool for assessing functional literacy in the health care setting. The TOFHLA's face validity is apparent given the fact that the material was drawn from commonly used hospital texts. Construct validity was shown by demonstrating statistically significant correlations with the REALM and the WRAT-R. After examining the performance of 200 English-speaking and 203 Spanish-speaking patients in two public hospital settings, it is apparent that the testing is also relevant given the high rates of reading comprehension problems found during test development. Fifty percent of the English-speaking and 74% of the Spanish-speaking patients had significant difficulty or were unable to read and interpret real health texts, and only 25% of the English-speaking and 9% of the Spanish-speaking patients could read and interpret most health texts contained in the TOFHLA.

Limited sampling for this test development study does not allow us to draw conclusions about the true prevalence and associations of low functional health literacy. Functional health literacy may be confounded by patients' overall familiarity with the health care system. For example, patients who have chronic medical conditions and worse functional status may have higher functional health literacy than would be expected based on their educational levels. In addition, testing in a larger population of patients is needed to determine appropriate cutoffs for scoring low, marginal, and adequate functional health literacy. This research is now in progress.

Further studies are also needed to determine whether an abbreviated form of the TOFHLA could be developed for brief screening in the clinical setting. Currently, most physicians probably just assume all their patients are functionally literate, even though nationally publicized literacy surveys document this is not the case. Screening tools are needed to identify low-literacy patients who may need special assistance to understand their diseases and comply with recommended treatments. The TOFHLA requires up to 22 minutes to administer, making it more useful as a research tool than a clinical tool, at this point. The Robert Wood Johnson Foundation "Literacy in Health

Care" project was begun in 1992 to address the relationship of literacy to health care. The TOFHLA appears to be an appropriate tool for measuring functional health literacy and should provide better insight into the problems that low-literacy patients face in the health care setting. This test development study suggests that low functional health literacy is prevalent among English- and Spanish-speaking patients at two urban public hospitals, and demonstrates that many patients cannot perform the basic tasks required of them.

The prevalence of low literacy in our society has tremendous importance for the delivery of health care, and it may be an important nonfinancial barrier to receiving high-quality care. Further investigation is required to assess not only the overall prevalence of low literacy, but also how it actually affects patients' abilities to understand their medical conditions and adhere to treatment recommendations.

The authors express their sincere thanks to Ron Hall, Project Coordinator, and Nina Parikh, Research Associate, for the Literacy in Health Care Project.

REFERENCES

1. Kirsch IS, Jungeblut A, Jenkins L, Kolstad A. Adult Literacy in America: A First Look at the Results of the National Adult Literacy Survey. Washington, DC: National Center for Education Statistics, U.S. Department of Education, 1993;xiv.
2. General Accounting Office. Report to Congress. The Adult Basic Education Program: Progress in Reducing Illiteracy and Improvement Needs. Washington, DC: U.S. Office of Education, 1975.
3. U.S. Department of Commerce, Bureau of the Census. English Language Proficiency Study. Washington, DC: Government Printing Office, 1982.
4. Davis TC, Crouch MA, Wills G, Miller S, Abdehou DM. The gap between patient reading comprehension and the readability of patient education materials. J Fam Pract. 1990;31:533–8.
5. Jolly BT, Scott JL, Feied CF, Sanford SM. Functional illiteracy among emergency department patients: a preliminary study. Ann Emerg Med. 1993;22:573–8.
6. Jackson RH, Davis TC, Bairnsfather LE, George RB, Crouch MA, Gault H. Patient reading ability: an overlooked problem in health care. South Med J. 1991;84:1172–5.
7. Davis TC, Crouch MA, Long SW, et al. Rapid assessment of literacy levels of adult primary care patients. Fam Med. 1991;23:433–5.
8. Weiss BD, Hart G, McGee DL, D'Estelle S. Health status of illiterate adults: relation between literacy and health status among persons with low literacy skills. J Am Board Fam Pract. 1992;5:257–64.
9. Jastak S, Wilkinson GS. Wide Range Achievement Test–Revised (WRAT-R). San Antonio, TX: The Psychological Corporation, 1984.
10. Davis TC, Long SW, Jackson RH, et al. Rapid estimate of adult literacy in medicine: a shortened screening instrument. Fam Med. 1993;25:391–5.
11. Nurss JR, Baker DW, Davis TC, Parker RM, Williams MV. Difficulties in functional health literacy screening in Spanish-speaking adults. J Reading. 1995;38:632–7.
12. Taylor WS. Cloze procedure: a new test for measuring readability. Journalism Q. 1953;30:415–33.
13. Laubach RS, Koschnick K. Using Readability: Formulas for Easy Adult Materials. Syracuse, NY: New Readers Press, 1977.
14. SPSS-X User's Guide. 3rd edition. Chicago, IL: SPSS, Inc., 1988.
15. Doak L, Doak C. Patient comprehension profiles: recent findings and strategies. Patient Couns Health Educ. 1980;2:101–6.