

# Quick Assessment of Literacy in Primary Care: The Newest Vital Sign

Barry D. Weiss, MD<sup>1</sup>

Mary Z. Mays, PhD<sup>2</sup>

William Martz, MD<sup>1</sup>

Kelley Merriam Castro, MA<sup>1</sup>

Darren A. DeWalt, MD, MPH<sup>3</sup>

Michael P. Pignone, MD, MPH<sup>3</sup>

Joy Mockbee, MD, MPH<sup>1</sup>

Frank A. Hale, PhD<sup>1</sup>

<sup>1</sup>University of Arizona College of Medicine, Department of Family and Community Medicine, Tucson, Ariz

<sup>2</sup>College of Nursing, Arizona State University, Tempe, Ariz

<sup>3</sup>Department of Internal Medicine, University of North Carolina, Chapel Hill, Chapel Hill, NC



MORE ONLINE  
www.annfammed.org

*Conflicts of interest: Drs Weiss, DeWalt, and Pignone have received research support and honoraria from the Pfizer's Health Literacy/Clear Health Communication Initiative. None of these individuals have any other financial interest in the Initiative or in Pfizer, Inc.*

## CORRESPONDING AUTHOR

Barry D. Weiss, MD

Department of Family and Community Medicine

University of Arizona College of Medicine

1450 North Cherry Ave

Tucson, AZ 85719

bdweiss@u.arizona.edu

## ABSTRACT

**PURPOSE** Current health literacy screening instruments for health care settings are either too long for routine use or available only in English. Our objective was to develop a quick and accurate screening test for limited literacy available in English and Spanish.

**METHODS** We administered candidate items for the new instrument and also the Test of Functional Health Literacy in Adults (TOFHLA) to English-speaking and Spanish-speaking primary care patients. We measured internal consistency with Cronbach's  $\alpha$  and assessed criterion validity by measuring correlations with TOFHLA scores. Using TOFHLA scores < 75 to define limited literacy, we plotted receiver-operating characteristics (ROC) curves and calculated likelihood ratios for cutoff scores on the new instrument.

**RESULTS** The final instrument, the Newest Vital Sign (NVS), is a nutrition label that is accompanied by 6 questions and requires 3 minutes for administration. It is reliable (Cronbach  $\alpha$  > 0.76 in English and 0.69 in Spanish) and correlates with the TOFHLA. Area under the ROC curve is 0.88 for English and 0.72 for Spanish versions. Patients with more than 4 correct responses are unlikely to have low literacy, whereas fewer than 4 correct answers indicate the possibility of limited literacy.

**CONCLUSION** NVS is suitable for use as a quick screening test for limited literacy in primary health care settings.

*Ann Fam Med* 2005;3:514-522. DOI: 10.1370/afm.405.

## INTRODUCTION

In 2004, the Institute of Medicine (IOM), the Agency for Healthcare Research and Quality (AHRQ), and the American Medical Association (AMA), all issued reports on health literacy.<sup>1-3</sup> Health literacy is the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions.<sup>4</sup> It involves the ability to use and interpret text, documents, and numbers effectively—skills that might seem to be distinct but are highly correlated with one another.<sup>1,5-7</sup>

The IOM, AHRQ, and AMA reports all noted that large segments of the American population—as many as one half of all adults—lack the literacy skills needed to function adequately in a health care environment. They would not, for example, be able to reliably and consistently determine the proper dose of cold medicine for a child, nor would they be able to read and understand informed consent documents.<sup>8,9</sup> Individuals with limited literacy come from all segments of society, and most are white, native-born Americans.<sup>10</sup>

Individuals with limited literacy have less knowledge about their health problems,<sup>11-16</sup> more hospitalizations,<sup>17,18</sup> higher health care costs,<sup>19,20</sup> and poorer health status<sup>21-25</sup> than those with adequate literacy. The relation between limited literacy and these factors is consistent across studies and persists after adjusting for confounding sociodemographic variables. With

awareness of patients' literacy skills, health information can be tailored for delivery to patients in a format they can understand.<sup>26</sup>

Although health literacy is a complex and multifaceted construct, researchers have developed instruments that assess literacy skills using health-context materials. Two such literacy assessments are widely used. One is the Test of Functional Health Literacy in Adults (TOFHLA),<sup>27,28</sup> which is the instrument most often used for literacy assessment in health care research. The TOFHLA is available in English and Spanish and has good psychometric characteristics, but the length of time required for administration of the TOFHLA (18 to 22 minutes for the full version and 7 to 10 minutes for a short version) precludes its use in busy primary care settings.<sup>29</sup> The second test, the Rapid Estimate of Adult Literacy in Medicine (REALM), can be administered quickly (less than 3 minutes) but it, too, has limitations. In particular, the REALM is only available in English.<sup>5,28</sup> This report describes the validation of a new rapid literacy assessment instrument in both English and Spanish, using the TOFHLA as the reference standard.

## METHODS

### Overview

We recruited English-speaking and Spanish-speaking patients from primary care clinics. The full TOFHLA and candidate test items for the new health literacy instrument were administered to the patients. Statistical tests were then used to determine which of the candidate test items from the new instrument best correlated with results of the TOFHLA. The University of Arizona Human Subjects Protection Committee approved our methods, and all participants gave informed consent.

### Instruments

#### Test of Functional Health Literacy in Adults

The TOFHLA is a 2-part test that is available in both English and Spanish. The first part provides participants with medical information or instructions about various scenarios, such as instructions on a prescription label or instructions about preparation for a diagnostic procedure. Participants review the scenarios and then answer questions that test their understanding of the information in the scenarios. The second part of the TOFHLA is based on the Cloze method, in which participants are given passages of text about medical topics with selected words deleted and replaced with blank spaces. The participants must fill in the blank spaces using words selected from a multiple choice list of options, identifying the words most appropriate to the context of the passage.

TOFHLA scores can range from 0 to 100, with

higher scores indicating better literacy. Scores of <60 represent inadequate literacy, 60 to 74 represent marginal literacy, and >75 represent adequate literacy. Individuals with TOFHLA scores in the inadequate or marginal range (ie, score of <75) would likely have trouble understanding written material that requires a 7th-grade reading level or higher, and often need assistance to understand completely instructions for their medical care.

Scores on the TOFHLA correlate with scores on standardized reading tests used in general education, such as the Revised Wide Range Achievement Test ( $r = 0.74$ )<sup>27,32</sup> and also with scores on the REALM ( $r = 0.84$ ).<sup>27</sup>

### Newest Vital Sign

Our new literacy screening tool, the Newest Vital Sign (NVS), was developed from a series of scenarios. Patients were given health-related information, which the patients read and then demonstrated their ability to use the information by answering questions about the scenarios. The questions were scored as either correct or incorrect according to a scoring key provided to the interviewers. The Spanish version was developed by translation and back-translation of the English version.

The development of the new instrument involved serially testing candidate scenarios and candidate questions on more than 1,000 patients. This preliminary testing was undertaken with the same patient population (though different patients) on which final testing was performed.

The original scenarios were developed by a panel of health literacy experts based on the concepts and types of scenarios used in health literacy research and in general literacy (reading and writing) assessments, such as the National Adult Literacy Survey<sup>33</sup> and the Instrument for Diagnosis of Reading,<sup>34</sup> and in health-literacy assessment instruments, such as the TOFHLA.

The scenarios and questions were refined after feedback from patients, interviewers, and data analysts about the clarity and ease of scoring of items. Ultimately, they were reduced to 5 candidate scenarios: (1) instructions from a prescription for headache medication, (2) a consent form for coronary angiography, (3) heart failure self-care instructions, (4) a nutrition label from an ice cream container, and (5) instructions for asthma medication that included a tapering steroid dose. These candidate scenarios varied in the type of literacy skills needed for understanding. Some, such as the angiography consent form and heart failure self-care instructions, emphasized the ability to understand text. Others, such as the nutrition label, placed greater emphasis on the ability to use numbers and mathematical concepts (numeracy). The inclusion of scenarios that involved both reading and numeracy skills was

driven by research indicating that these skills are highly correlated with one another,<sup>1,5-7</sup> plus an intuitive understanding that patients must be able to use and understand both text and numbers if they are to successfully deal with today's health care system.

A total of 21 questions (3 to 6 questions per scenario) accompanied the 5 scenarios. This 21-item pool of questions was administered to 500 participants in this study.

A final short form of the test, the NVS, was selected from this 21-item pool on the basis of its psychometric properties. The NVS uses 1 scenario (the ice cream nutrition label) and 6 questions (Figures 1a and 1b, and Supplemental Appendix 1, which is available online only at: <http://www.annfammed.org/cgi/content/full/3/6/514/DC1>).

The reliability, validity, and accuracy of the 1-scenario 6-item final version are reported in detail in this report. A summary of the reliability and validity of the other scenarios is provided in Table 1.

### Participants

Participants were patients from 3 primary care practices in Tucson, Ariz, all of which are affiliated with the University of Arizona College of Medicine. Two practices are operated by the college's faculty practice plan, and 1 is a publicly funded clinic located in a primarily Spanish-speaking area of Tucson.

### Sample Size

Based on pilot data indicating that the prevalence of adequate health literacy in our patient population is about 75% (defined as a score >74 on the TOFHLA), sampling error calculations revealed that with a random sample of 492 patients, we could be 95% confident that 71.5% to 78.5% of the participants would score >74 on the TOFHLA. Thus, we collected data on 500 patients (250 in each language group). This sample size provided a power of 0.90 for an independent samples *t* test, when mean differences between groups were equal to 0.40 standard deviations and  $\alpha$  was set at 0.05.

### Participant Recruitment

Participants had to be 18 years old or older, speak English or Spanish, have visual acuity sufficient to read the instruments being tested, and have grossly normal cognitive function that was adequate to interact with study personnel. Participants completing the interview received a \$20 supermarket gift certificate.

Bilingual project staff interviewers approached consecutive patients in the waiting rooms of the 3 clinics during specified periods. If the patient agreed and met eligibility criteria, the interviewer administered both the TOFHLA and the candidate items for the NVS. We

Figure 1A. The newest vital sign — English.

<b>Nutrition Facts</b>	
Serving Size	1/2 cup
Servings per container	4
<b>Amount per serving</b>	
Calories 250	Fat Cal 120
	%DV
<b>Total Fat 13g</b>	20%
Sat Fat 9g	40%
<b>Cholesterol 28mg</b>	12%
<b>Sodium 55mg</b>	2%
<b>Total Carbohydrate 30g</b>	12%
Dietary Fiber 2g	
Sugars 23g	
<b>Protein 4g</b>	8%

\* Percent Daily Values (DV) are based on a 2,000 calorie diet. Your daily values may be higher or lower depending on your calorie needs.

**Ingredients:** Cream, Skim Milk, Liquid Sugar, Water, Egg Yolks, Brown Sugar, Milkfat, Peanut Oil, Sugar, Butter, Salt, Carrageenan, Vanilla Extract.

Note: This single scenario is the final English version of the newest vital sign. The type size should be 14-point (as shown above) or larger. Patients are presented with the above scenario and asked the questions shown in Figure 1b.

**Figure 1B. Questions and answers score sheet for the newest vital sign — English.**

	ANSWER CORRECT?	
	YES	NO
<b>READ TO SUBJECT:</b> This information is on the back of a container of a pint of ice cream.		
<b>QUESTIONS</b>		
1. If you eat the entire container, how many calories will you eat? <b>Answer</b> <input type="checkbox"/> 1,000 is the only correct answer	_____	_____
2. If you are allowed to eat 60 g of carbohydrates as a snack, how much ice cream could you have? <b>Answer</b> Any of the following is correct: <input type="checkbox"/> 1 cup (or any amount up to 1 cup) <input type="checkbox"/> Half the container Note: If patient answers "2 servings," ask "How much ice cream would that be if you were to measure it into a bowl?"	_____	_____
3. Your doctor advises you to reduce the amount of saturated fat in your diet. You usually have 42 g of saturated fat each day, which includes 1 serving of ice cream. If you stop eating ice cream, how many grams of saturated fat would you be consuming each day? <b>Answer</b> 33 is the only correct answer	_____	_____
4. If you usually eat 2500 calories in a day, what percentage of your daily value of calories will you be eating if you eat one serving? <b>Answer</b> 10% is the only correct answer	_____	_____
Pretend that you are allergic to the following substances: Penicillin, peanuts, latex gloves, and bee stings.		
5. Is it safe for you to eat this ice cream? <b>Answer</b> <input type="checkbox"/> No	_____	_____
6. (Ask only if the patient responds "no" to question 5): Why not? <b>Answer</b> Because it has peanut oil.	_____	_____
<b>Total Correct</b>	_____	_____

gave patients a hard copy of the NVS nutrition label to hold and to which they could refer, as needed, while the interviewer asked the 6 questions out loud. Patients were testing in the language they preferred for reading.

This process continued until 250 English-speaking and 250 Spanish-speaking patients had been tested. The order of administering the tests was alternated so that interviewers administered the NVS questions first to even-numbered participants and the TOFHLA first to odd-numbered participants.

### Data Analysis

We used means, standard deviation (SD), standard error of the means, histograms, *t* tests, and analysis of covariance to summarize the participant demographic characteristics and their performance on the tests. The TOFHLA was scored according to the instructions provided with the instrument. Candidate items on the

NVS test were scored by giving 1 point for each correct answer.

Analysis of the psychometric properties of the English and Spanish versions of the NVS test and the TOFHLA were conducted separately using identical methods. Reliability of the NVS was assessed in terms of internal consistency (Cronbach  $\alpha$ ). Criterion validity was determined by calculating the correlation (Pearson *r*) between scores on the NVS and TOFHLA. We quantified the relative accuracy of age, educational level, and NVS scores as predictors of adequate literacy (defined by TOFHLA scores >74) by computing their receiver operating characteristic (ROC) curves. The ROC curves were used to calculate the sensitivity and specificity for selected cutoff scores on the NVS test. Stratum-specific likelihood ratios were calculated for each NVS score.

### RESULTS

We enrolled 250 English-speaking participants to validate the English version of the NVS (NVS-E) and 250 Spanish-speaking participants to validate the Spanish version (NVS-S). These 500 participants represented approximately 80% of the individuals asked to participate

in the study. Participants' demographic characteristics are shown in Table 2.

Table 2 also provides a comparison of the NVS with TOFHLA scores of the English and Spanish samples before and after adjustment for sex and educational level. The English sample had significantly higher scores than the Spanish sample on both instruments; these differences could not be explained by differences in sex and educational level.

### Newest Vital Sign: English

The number of correct items on the NVS-E ranged from 0 to 6 (mean = 3.4  $\pm$  1.9). The time required to administer the 6 items was recorded for a series of 24 participants; the average time was 2.9 minutes (SD 1.2 minutes; range = 1.5-6.2 minutes). Total scores on the English version of the TOFHLA (TOFHLA-E) ranged from 12 to 100 (mean = 86.3  $\pm$  14.1). As shown in

**Table 1. Reliability and Validity of English and Spanish Candidate Scenarios and Total Test**

Scenario	Number of Items	Reliability: Cronbach $\alpha$	Validity*: Pearson $r$
English			
1. Prescription for headache medication	3	0.23	0.43
2. Consent form for angiography	4	0.40	0.51
3. Self-care instructions for heart failure	5	0.38	0.20
4. Nutrition label from ice cream <sup>†</sup>	6	0.76	0.59
5. Instructions for tapering prednisone	3	0.66	0.35
Total	21	0.78	0.64
Spanish			
1. Prescription for headache medication	3	0.37	0.38
2. Consent form for angiography	4	0.37	0.55
3. Self-care instructions for heart failure	5	0.33	-0.01
4. Nutrition label from ice cream <sup>†</sup>	6	0.69	0.49
5. Instructions for tapering prednisone	3	0.39	0.37
Total	21	0.69	0.59

\* Correlation of scenario score and score on criterion (ie, Test of Functional Health Literacy in Adults, Spanish Version).

<sup>†</sup> This scenario (nutrition label) was selected as the single best scenario, and is the content of the Newest Vital Sign – Spanish.

**Table 2. Demographic Characteristics and Scores on the English and Spanish Language Versions of the TOFHLA and NVS**

Characteristics and Scores	English (n = 250)	Spanish (n = 250)
<b>Demographics</b>		
Age (years)		
Mean (SD)	41.3 (14.8)	40.8 (13.9)
Range	18-85	18-77
Ethnic/racial group (%)		
Non-Hispanic white	43	0
Hispanic	43	100
African American	5	0
Native American	4	0
Unknown or other	5	0
Education (y)		
Mean (SD)	12.7 (2.8)	10.7 (3.2)
Range	2-24	0-23
Insurance status (%)		
Private	66	45
Government programs	27	32
Uninsured	7	22
<b>Test scores</b>		
TOFHLA, mean (SEM)*	86.3 (0.9)	75.7 (1.2)
NVS – mean (SEM) <sup>†</sup>	3.4 (1.9)	1.6 (1.5)
TOFHLA adjusted, mean (SEM) <sup>‡</sup>	84.6 (1.0)	77.4 (1.0)
NVS adjusted, mean (SEM) <sup>§</sup>	33 (0.1)	1.8 (0.1)

TOFHLA = Test of Functional Health Literacy in Adults; NVS = Newest Vital Sign; SEM = standard error of the mean.

\* Comparison by independent samples *t* test between English and Spanish samples,  $P < .001$ .

<sup>†</sup> Comparison by independent samples *t* test between English and Spanish samples,  $P < .001$ .

<sup>‡</sup> Comparison by analysis of covariance between English and Spanish samples, adjusted for sex and years of education,  $P < .001$ .

<sup>§</sup> Comparison by analysis of covariance between English and Spanish samples, adjusted for sex and years of education,  $P < .001$ .

Figure 2, the distribution of TOFHLA-E scores was severely negatively skewed.

There was no significant difference in scores between men (mean =  $3.3 \pm 2.0$ ;  $n = 55$ ) and women (mean =  $3.5 \pm 1.9$ ;  $n = 195$ ) on the NVS-E ( $P > .05$ ). In contrast, there was significant sex bias on the TOFHLA-E, with men scoring significantly lower (mean =  $82.5 \pm 17.1$ ;  $n = 55$ ) than women (mean =  $87.4 \pm 13.0$ ;  $n = 195$ ;  $P < .001$ ). This bias is not the result of confounding with age or educational level, as there were no significant differences between men and women in either of these variables ( $P > .05$ ).

#### Reliability, Validity, and Accuracy

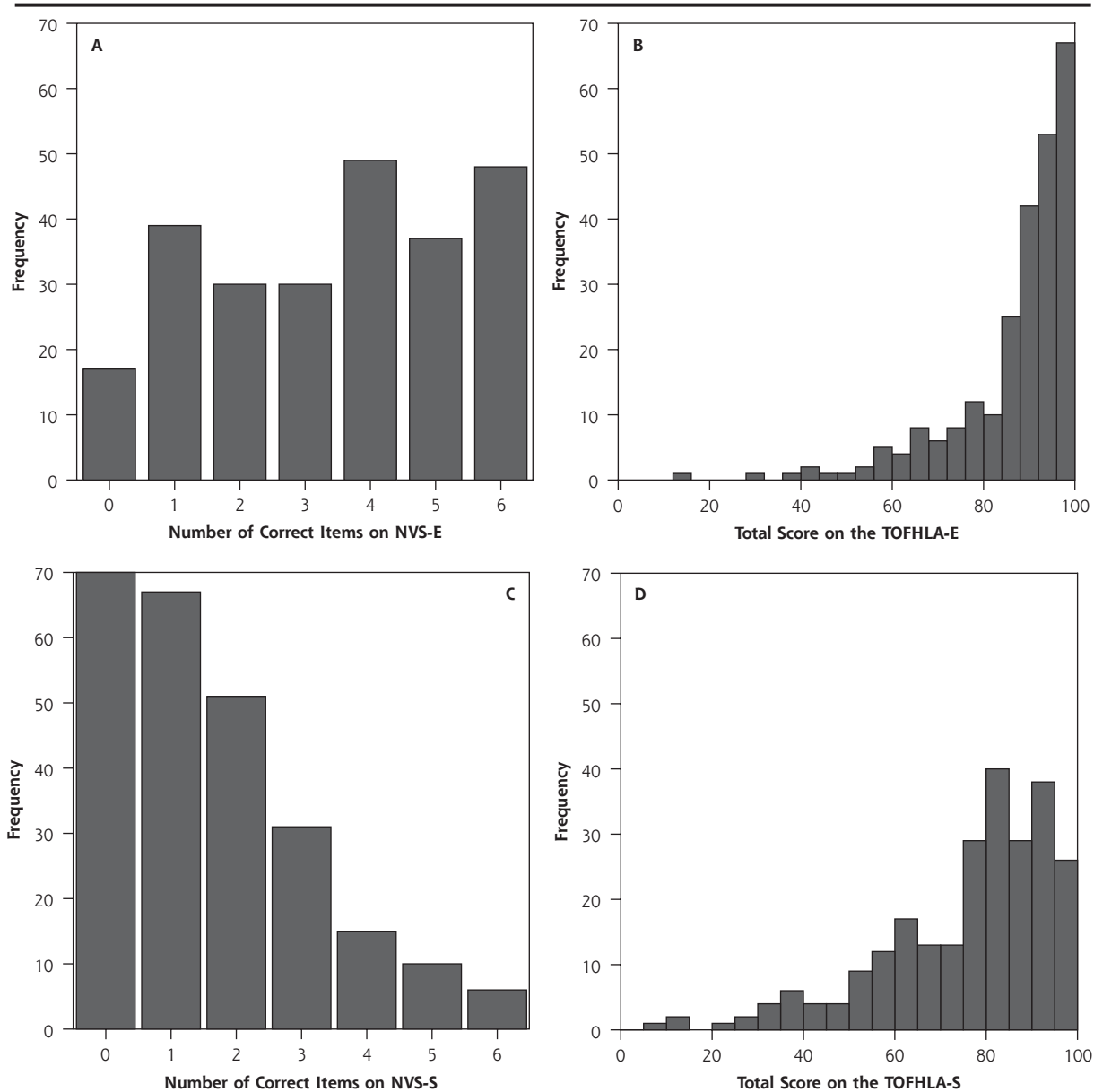
The internal consistency of the NVS-E was good (Cronbach  $\alpha = 0.76$ ), as was the criterion validity ( $r = 0.59$ ,  $P < .001$ ). Supplemental Appendix 2 (available online only at <http://www.annfammed.org/cgi/content/full/3/6/514/DC1>) plots the relationship between scores on the NVS-E and the TOFHLA-E.

The area under the ROC curve for predicting TOFHLA-E scores was 0.88 (95% CI, 0.84-0.93;  $P < .001$ ) for the NVS-E, substantially higher than the 0.72 (95% CI, 0.63-0.81;  $P < .001$ ) found for educational level or the 0.71 (95% CI, 0.63-0.79;  $P < .001$ ) found for age. Thus, the NVS-E score was more accurate than educational level or age.

The ROC curve for the NVS-E showed that a score of  $< 2$  on the NVS-E had a sensitivity of 72% and specificity of 87% for predicting limited literacy (TOFHLA score  $< 75$ ), whereas a score of  $< 4$  had a sensitivity of 100% and a specificity of 64%. Stratum-specific likelihood ratios for cutoff scores on the NVS-E are shown in Table 3 under English.



**Figure 2. Distribution of scores on the Newest Vital Signs (NVS) and the Test of Functional Health Literacy in Adults (TOFHLA).**



Note: Histograms on top show the distribution of participants' scores on (A) the English version of the Newest Vital Sign (NVS-E) and (B) the English version of the Test of Functional Health Literacy in Adults (TOFHLA-E). Histograms on the bottom show the distribution of scores on (C) the Spanish version of the Newest Vital Sign (NVS-S) and (D) the Spanish version of the Test of Functional Health Literacy in Adults (TOFHLA-S).

These ratios show, for example, that getting only 1 item correct on the NVS-E has a stratum-specific likelihood ratio of 5.1 for marginal or inadequate literacy, a 4-fold increase over that seen with getting 2 items correct.

### Newest Vital Sign: Spanish

The number of correct items on the NVS-S ranged from 0 to 6 (mean =  $1.6 \pm 1.5$ ). The time required to administer the 6 items was recorded for a series of 36 participants;

the average time was 3.4 minutes (SD 1.2 minutes; range = 2.1- 8.2 minutes). Total scores on the Spanish version of the TOFHLA (TOFHLA-S) ranged from 8 to 100 (mean =  $75.7 \pm 18.5$ ). As shown in Figures 2C and 2D, the distribution of NVS-S scores was quite different from that of TOFHLA-S scores; NVS-S scores were positively skewed while TOFHLA-S scores were negatively skewed.

There were no significant differences ( $P > .05$ ) between men ( $n = 29$ ) and women ( $n = 221$ ) on the

NVS-S (mean =  $1.5 \pm 1.5$ , and mean =  $1.6 \pm 1.5$ , respectively) or on the TOFHLA-S (mean =  $71.0 \pm 18.2$ , and mean =  $76.3 \pm 18.5$ , respectively). Nor were there any significant differences between men and women in age or educational level ( $P > .05$ ).

### Reliability, Validity, and Accuracy

The internal consistency of the NVS-S was good (Cronbach  $\alpha = 0.69$ ), as was the correlation with the TOFHLA ( $r = 0.49$ ,  $P < .001$ ). Supplemental Appendix 2 plots the relationship between scores on the NVS-S and the TOFHLA-S.

The area under the ROC curve for predicting TOFHLA-S scores was 0.72 (95% CI, 0.66-0.79;  $P < .001$ ) for the NVS-S, slightly higher than the 0.69 (95% CI, 0.62-0.76;  $P < .001$ ) found for educational level or the 0.64 (95% CI, 0.56-0.71;  $P < .001$ ) found for age. The ROC curve for the NVS-S showed that scoring  $< 2$  on the NVS-S had a sensitivity of 77% and a specificity of 57% for predicting limited literacy (TOFHLA-S score  $< 75$ ), whereas scoring  $< 4$  had a sensitivity of 100% and a specificity of 19%. Stratum-specific likelihood ratios for cutoff scores on the NVS-S are shown in Table 3 under Spanish.

## DISCUSSION

In samples of English-speaking and Spanish-speaking primary care patients in the southwestern United States, the NVS—a 6-item test based on the ability to read and apply information from a nutrition label—was a reliable and accurate measure of literacy with high sensitivity for detecting persons with limited literacy. The NVS is the first literacy screening test available in both English and Spanish that can be administered in approximately 3 minutes. It will permit clinicians or health care administrators to rapidly assess literacy in their patients.

### Properties and Clinical Significance of the NVS

The NVS had good sensitivity; in fact, based on the distribution of scores (Figure 2), NVS may be more sensitive than the TOFHLA to marginal health literacy. Its specificity, although less than optimal, is similar to or better than that of other widely used clinical screening methods, such as questionnaires to detect alcohol abuse,<sup>35,36</sup> breast self-examinations to screen for cancer,<sup>37</sup> and methods to detect arthritis<sup>38</sup> and measure osteoporosis risk.<sup>39</sup> Although the specificity of NVS may result in overestimating the percentage of patients with limited literacy, using the test can alert physicians

**Table 3. Stratum-Specific Likelihood Ratios for Cutoff Scores on the English and Spanish Versions of the Newest Vital Sign, Stratified by TOFHLA Score**

Number Correct on NVS-E	TOFHLA $< 75^*$	TOFHLA $\geq 75^*$	SSLR <sup>†</sup>	SSLR <sup>‡</sup> Grouped
<b>English</b>				
0	9	8	6.1	} 5.4
1	19	20	5.1	
2	5	25	1.1	} 1.2
3	6	24	1.4	
4	0	49	0	} 0
5	0	37	0	
6	0	48	0	
Total	39	211		
<b>Spanish</b>				
0	38	32	2.2	} 1.8
1	30	37	1.5	
2	17	34	0.9	} 0.6
3	3	28	0.2	
4	0	15	0	} 0
5	0	10	0	
6	0	6	0	
Total	88	162		

NVS-E = Newest Vital Sign-English; TOFHLA = Test of Functional Health Literacy in Adults  
SSLR = stratum-specific likelihood ratio.

\* 75 is the TOFHLA cut-off score below which individuals are considered to have marginal or inadequate literacy. Individuals with scores  $\geq 75$  are considered to have adequate literacy.

† Stratified by TOFHLA scores  $< 75$  (marginal or inadequate literacy) vs  $\geq 75$  (adequate literacy).

‡ SSLRs for detecting low literacy at each score on the NVS in English (1-A) and Spanish (1-B). Because of the similarity in SSLRs in adjacent groups, we suggest grouping scores into 3 categories: 0-1, 2-3, 4-6.

to patients who may need more attention and help physicians focus on physician-patient communication using recommended techniques.<sup>26,40</sup>

All patients who score  $> 4$  on the NVS will have adequate literacy when measured by the TOFHLA. A score  $< 4$  on the NVS, on the other hand, indicates the possibility of limited literacy. Clinicians should be particularly careful in their communication with patients who score  $< 2$ , as they have a greater than 50% chance of having marginal or inadequate literacy skills. Such patients cannot be reliably identified by questions about their education level, as education does not always predict of literacy—it only measures the number of years an individual attended school. Indeed, about one quarter of participants who scored at the very lowest of 5 literacy levels on the US Department of Education's National Adult Literacy Survey were high school graduates.<sup>33</sup>

### The Role of Numeracy in Health Literacy

Of the 5 candidate scenarios evaluated in this validation study, the one most effective at discriminating

low literacy from adequate literacy was the scenario requiring the most complex numeracy skills (Table 1). As discussed earlier, our testing included scenarios that assessed both reading and numeracy skills, but in every analysis we performed, the nutrition label with its quantitative-numerical questions was the best predictor of literacy (using the TOFHLA as a reference standard). That the questions from the nutrition label scenario had such high internal consistency suggests answering those questions involved not only a math skill, but also a locate-the-information skill (by reading and comprehending) and an abstract reasoning skill (eg, imagining they have an allergy to peanuts, noting that even vanilla ice cream can have a peanut product in it, and reasoning that peanut oil is probably not good for you if you are allergic to peanuts).

Furthermore, the use of a nutrition label to assess health literacy is intuitively appealing because nutrition labels are familiar items that are important parts of health management for many chronic diseases. They are also used for health promotion in that many healthy people use information on nutrition labels to help achieve health eating habits. Patients' ability to understand and use the information on nutrition labels has been the subject of study in a host of health promotion and epidemiology research projects, both in the United States and elsewhere.<sup>41-43</sup>

### Limitations

We compared the NVS to the full version of the TOFHLA, not the short-version TOFHLA that is in more widespread use. The full TOFHLA, however, is the standardized instrument from which the short version was derived, so its psychometric properties are an appropriate reference standard for the development of new instruments.

Health literacy is a complex construct that encompasses many aspects of how individuals use health information and the health care system. Our test, like the TOFHLA and the REALM, measures reading and interpretation skills (ie, general literacy, reasoning, and the ability to use numbers) as applied to material with health content, rather than all aspects of health literacy.<sup>32,44-45</sup>

The psychometric properties of the Spanish version of the NVS, although adequate to screen patients for limited literacy, were not as good as those of the English version. This fact may stem from the greater heterogeneity of language and culture among our Spanish-speaking patients, who come from all regions of South America, Central America, and Mexico.

Finally, the primary care practices in this study were selected because of their high percentage of Spanish-speaking patients, and among the Spanish-speaking participants the percentage of male patients

was relatively small. These practices do not, therefore, have demographics that are fully representative of all primary care practices in the United States. Testing of the NVS on other patient populations could further validate the accuracy of the instrument.

The NVS has advantages over currently available instruments. Specifically, it is available in Spanish, whereas the REALM is not, and it can be administered much more quickly than the TOFHLA. The NVS also does not have the ceiling effect seen with the TOFHLA and, therefore, particularly in the English version, the NVS provides better discrimination of skill levels among individuals in the upper part of the distribution of literacy skills. Future investigations should examine (1) how to best introduce and implement NVS in primary care practice, (2) the validity of NVS in other primary care practices and also in non-primary care settings, (3) whether raising clinicians' awareness of patients' literacy by using NVS results in improved clinician-patient communication and better health outcomes, and (4) whether a similar nutrition label scenario can assess literacy in speakers of languages other than English and Spanish.

**To read or post commentaries in response to this article, see it online at <http://www.annfammed.org/cgi/content/full/3/6/514>.**

**Key words:** Literacy; health literacy; physician-patient communication; Spanish

Submitted April 30, 2005; submitted, revised, July 26, 2005; accepted July 31, 2005.

A version of this report was presented at Medical Futures Forum, New York, NY, March 24, 2005.

**Funding support:** This study was supported by funding from the Pfizer Clear Health Communication Initiative.

**Acknowledgments:** The authors thank Barbara DeBuono, MD, MPH, and Lisa Dieter of Pfizer, Inc, and Ana Rita Gonzales, ScD, of Fleishman-Hillard for their support of this project. The authors also acknowledge and appreciate the assistance of John Blackburn, Katie O'Brien, and Berenis Romo, who conducted health literacy testing. Additional thanks to Maria Chavez, research specialist in the University of Arizona Department of Family and Community Medicine, for her help with data entry.

### References

1. Nielson-Bohlman L, Panzer A, Kindig D, eds. *Health Literacy: A Prescription to End Confusion*. Washington, DC: National Academies Press; 2004.
2. Berkman N, DeWalt D, Pignone M, et al. Literacy and Health Outcomes. Summary, Evidence Report/Technology Assessment No. 87 (Prepared by RTI International - University of North Carolina Evidence-Based Practice Center under Contract No. 290-02-0016). Rockville, Md: Agency for Healthcare Research and Quality; 2004. AHRQ Publication No. 04-E007-1.
3. Schwartzberg J, VanGeest J, Wang C, eds. *Understanding Health Literacy: Inspirations for Medicine and Public Health*. Chicago, Ill: American Medical Association Press; 2004.



4. US Department of Health and Human Services. Health Communication. In: *Healthy People 2010*. 2nd ed. Washington, DC: US Government Printing Office; 2000.
5. Gottfredson L. Why g matters: the complexity of everyday life. *Intelligence*. 1997;24:79-132.
6. Reder S. What Does NALS Measure? Issues of Dimensionality and Construct Validity. Portland, Ore: Portland State University, National Center on Adult Literacy; 1995.
7. Baldwin J, Kirsch L, Rock D, Yamamoto K. *The Literacy Proficiencies of GED Examinees: Results From the GED-NALS Comparison Study*. Washington, DC: American Council on Education and Educational Testing; 1995.
8. Rudd R, Kirsch L, Yamamoto K. *Literacy In Health Contexts: A Re-Analysis Of The NALS*. American Public Health Association Annual Conference Session 4182.1. Washington, DC, American Public Health Association; 2003.
9. Paasche-Orlow MK, Taylor HA, Brancati FL. Readability standards for informed-consent forms as compared with actual readability. *N Engl J Med*. 2003;348:721-726.
10. Weiss B. Epidemiology of low literacy. In: Schwartzberg J, VanGeest J, Wang C, eds. *Understanding Health Literacy: Implications for Medicine and Public Health*. Chicago, Ill: American Medical Association Press; 2004:17-42.
11. 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.
12. Arnold CL, Davis TC, Berkel HJ, et al. Smoking status, reading level, and knowledge of tobacco effects among low-income pregnant women. *Prev Med*. 2001;32:313-320.
13. Williams MV, Baker DW, Honig EG, Lee TM, Nowlan A. Inadequate literacy is a barrier to asthma knowledge and self-care. *Chest*. 1998;114:1008-1015.
14. Davis TC, Arnold C, Berkel HJ, et al. Knowledge and attitude on screening mammography among low-literate, low-income women. *Cancer*. 1996;78:1912-1920.
15. Davis TC, Dolan NC, Ferreira MR, et al. The role of inadequate health literacy skills in colorectal cancer screening. *Cancer Invest*. 2001;19:193-200.
16. Gazmararian JA, Williams MV, Peel J, Baker DW. Health literacy and knowledge of chronic disease. *Patient Educ Couns*. 2003;51:267-275.
17. Baker DW, Gazmararian JA, Williams MV, et al. Functional health literacy and the risk of hospital admission among Medicare managed care enrollees. *Am J Public Health*. 2002;92:1278-1283.
18. Baker DW, Parker RM, Williams MV, Clark WS. Health literacy and the risk of hospital admission. *J Gen Intern Med*. 1998;13:791-798.
19. Weiss BD, Palmer R. Relationship between health care costs and very low literacy skills in a medically needy and indigent Medicaid population. *J Am Board Fam Pract*. 2004;17:44-47.
20. Howard DH, Gazmararian J, Parker RM. The impact of low health literacy on the medical costs of Medicare managed care enrollees. *Am J Med*. 2005;118:371-377.
21. 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.
22. Kalichman SC, Rompa D. Functional health literacy is associated with health status and health-related knowledge in people living with HIV/AIDS. *J Acquir Immune Defic Syndr*. 2000;25:337-344.
23. Schillinger D, Grumbach K, Piette J, et al. Association of health literacy with diabetes outcomes. *JAMA*. 2002;288:475-482.
24. 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;5:257-264.
25. Bennett CL, Ferreira MR, Davis TC, et al. Relation between literacy, race, and stage of presentation among low-income patients with prostate cancer. *J Clin Oncol*. 1998;16:3101-3104.
26. Weiss B. *Health Literacy: A Manual for Clinicians*. Chicago, Ill: American Medical Association Foundation; 2003.
27. 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.
28. Baker DW, Williams MV, Parker RM, Gazmararian JA, Nurss J. Development of a brief test to measure functional health literacy. *Patient Educ Couns*. 1999;38:33-42.
29. Woodwell D, Cherry D. *National Ambulatory Medical Care Survey: 2002 Summary. Advanced Data from Vital and Health Statistics*. Hyattsville, Md: National Center for Health Statistics; 2004. No. 346.
30. Davis TC, Long SW, Jackson RH, et al. Rapid estimate of adult literacy in medicine: a shortened screening instrument. *Fam Med*. 1993;25:391-395.
31. Taylor W. Cloze procedure: a new tool for measuring readability. *Journalism Q*. 1953;30:415-433.
32. Jastak S, Wilkinson G. *Wide Range Achievement Test - Revised 3*. Wilmington, Del: Jastak Associates; 1993.
33. Kirsch I, 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. US Department of Education; 1993.
34. Blanchard J, Garcia H, Carter R. *Instrumento para Diagnostica Lecturas (Español-English): Instrument for the Diagnosis of Reading*. Dubuque, Ia: Kendall-Hunt Publishing Co; 1989.
35. Bradley KA, Boyd-Wickizer J, Powell SH, Burman ML. Alcohol screening questionnaires in women: a critical review. *JAMA*. 1998;280:166-171.
36. Cook RL, Chung T, Kelly TM, Clark DB. Alcohol screening in young persons attending a sexually transmitted disease clinic. Comparison of AUDIT, CRAFFT, and CAGE instruments. *J Gen Intern Med*. 2005;20:1-6.
37. Elmore JG, Armstrong K, Lehman CD, Fletcher SW. Screening for breast cancer. *JAMA*. 2005;293:1245-1256.
38. Bombard JM, Powell KE, Martin LM, Helmick CG, Wilson WH. Validity and reliability of self-reported arthritis: Georgia senior centers, 2000-2001. *Am J Prev Med*. 2005;28:251-258.
39. Mauck KF, Cuddihy MT, Atkinson EJ, Melton LJ, 3rd. Use of clinical prediction rules in detecting osteoporosis in a population-based sample of postmenopausal women. *Arch Intern Med*. 2005;165:530-536.
40. Doak C, Doak L, Root J. *Teaching Patients with Low Literacy Skills*. 2nd ed. Philadelphia, Pa: JB Lippincott Company; 1996.
41. Cowburn G, Stockley L. Consumer understanding and use of nutrition labeling: a systematic review. *Public Health Nutr*. 2005;8:21-28.
42. Pelletier AL, Chang WW, Delzell JE, Jr., McCall JW. Patients' understanding and use of snack food package nutrition labels. *J Am Board Fam Pract*. 2004;17:319-323.
43. Byrd-Bredbenner C, Cottee A. Consumer understanding of US and EU nutrition labels. *Br Food J*. 2000;102:615-629.
44. Markwardt F. *Peabody Individual Achievement Test - Revised*. Circle Pines, Minn: American Guidance Service; 1989.
45. Slosson R. *Slosson Oral Reading Tests - Revised*. East Aurora, NY: Slosson Educational Publishers; 1990.