The Health Literacy Skills Instrument: A 10-Item Short Form

Carla M. Bann [a] , Lauren A. McCormack [a] , Nancy D. Berkman [a] &
Linda B. Squiers [a]

[a] RTI International, Research Triangle Park, North Carolina, USA
Version of record first published: 03 Oct 2012.

PLEASE SCROLL DOWN FOR ARTICLE

Routledge
Taylor & Francis Group

The Health Literacy Skills Instrument: A 10-Item Short Form

CARLA M. BANN, LAUREN A. McCORMACK,
NANCY D. BERKMAN, AND LINDA B. SQUIERS

RTI International, Research Triangle Park, North Carolina, USA

The 25-item Health Literacy Skills Instrument (HLSI) was designed to measure the ability to read and understand text and locate and interpret information in documents (print literacy), to use quantitative information (numeracy), to listen effectively (oral literacy), and to seek information through the Internet (navigation). It is a publically available measure that can be used in surveillance activities, to evaluate interventions, and in research examining the relation between health literacy and health outcomes. The authors developed a 10-item, short form (SF) version of the HLSI, the HLSI-SF, using data gathered for the development of the longer form. The authors selected 10 items for inclusion in the HLSI-SF and conducted a confirmatory factor analysis and item response theory analyses, then computed Cronbach's alpha. The HLSI-SF demonstrated acceptable internal consistency reliability ($\alpha = .70$) for use in group-level comparisons. The HSLI-SF has many of the same advantages of the longer version with the additional benefit of taking only approximately 5 to 10 min to administer. The HLSI-SF offers researchers and practitioners a valid and reliable measure of health literacy skills.

The National Action Plan to Improve Health Literacy provides a plan for a multisector effort to improve health literacy and meet the objectives set forth in Healthy People 2020 (U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion, 2010). It calls for an increase in basic research and the development, implementation, and evaluation of practices and interventions to improve health literacy. To achieve this goal, the plan emphasizes the need to measure individual and population health literacy skills and to include health literacy measures in national and other surveys.

The most commonly used measures of health literacy in research are the Rapid Estimate of Adult Literacy in Medicine (REALM) and the Test of Functional Health Literacy in Adults (TOFHLA; Berkman et al., 2011). Although the National Assessment of Adult Literacy may have offered the most comprehensive list of health literacy–related questions, it has not been made available to researchers and, therefore, cannot be replicated in other studies (Weiss, 2009). Earlier instruments include the REALM (Davis et al., 1993), which is a word recognition and pronunciation test; the

TOFHLA (Parker, Baker, Williams, & Nurss, 1995), which requires subjects to fill in missing words in passages; Newest Vital Sign (Weiss et al., 2005), which is a documents and quantitative literacy skills tests; and the Schwartz and Woloshin Numeracy Test (Schwartz, Woloshin, Black, & Welch, 1997), which is limited to evaluating numeracy. The Department of Education's 2003 National Assessment of Adult Literacy survey measures navigation of the health care system as well as print literacy and numeracy (Kutner, Greenberg, Jin, & Paulsen, 2006).

Many of these instruments include a relatively large number of questions that may be time consuming to administer (REALM = 66 items, TOFHLA = 67 items, National Assessment of Adult Literacy = 28 items). Although more lengthy measures of health literacy skills may provide optimal validity and reliability, a shorter version enhances the feasibility of measuring health literacy in the clinical setting or as part of a large-scale surveillance instrument.

Shorter versions of the TOFHLA (S-TOFHLA) and REALM (REALM-R and REALM-SF; Arozullah et al., 2007; Bass, Wilson, & Griffith, 2003) have been developed. The nine-item S-TOFHLA is commonly used in research studies, whereas the shorter version of the REALM has been used less extensively (Arozullah et al., 2007; Berkman et al., 2011; Davis et al., 1993). More recently, the six-question Newest Vital Sign instrument was developed to quickly measure document literacy and numeracy aspects of health literacy in clinical settings (Weiss et al., 2005).

We developed the 25-item Health Literacy Skills Instrument (HLSI) as a skills-based tool to measure health literacy (McCormack et al., 2010). The skills include the ability to read and understand text and locate and interpret information in documents (print literacy), to use quantitative information (numeracy), to listen effectively (oral literacy), and to seek information through the Internet (navigation). Similar to other alternative shorter alternative measures of health literacy, the HLSI-Short Form (HLSI-SF) provides a shorter, validated alternative to its more extensive parent instrument. (The HLSI-SF and the 25-item HLSI measure a range of skills and are publicly available at http://www.rti.org/page.cfm/Health_Communication_and_ Marketing).

Method

Participants

We administered the health literacy questions using the KnowledgePanel, created by Knowledge Networks, an online nonvolunteer access panel. Potential panel members are chosen via a statistically valid sampling method and using known published sampling frames that cover 99% of the U.S. population. Address-based sampling, which is based on the U.S. Postal Service Delivery Sequence File, was used to select a probability sample of all U.S. households. This sample also comprises cell-phone households as well as non-Internet households. Address-based sampling is one of the most innovative means of obtaining nationally representative samples at minimum cost. Sampled non-Internet households are provided a laptop computer and free Internet service. KnowledgePanel consists of about 50,000 U.S. residents, aged 18 years or older, including persons of Hispanic origin that were selected probabilistically (for more information about the panel, see http://www.knowledgenetworks.com/ knpanel/index.html). Between October 7, 2009, and November 19, 2009, a total of 2,212 Knowledge Network panelists aged 18 or over were invited to participate in the survey.

Measures

The 25-item health literacy measure was developed through an extensive multistep scale development and evaluation process (McCormack et al., 2010) and is designed to address Ratzan and Parker's (2000) definition of *health literacy* (with minor modifications to wording) as "[t]he degree to which individuals can obtain, process, understand, and communicate about health-related information needed to make informed health decisions" (Berkman, Davis, & McCormack, 2010, p. 16). Cognitive interviews were conducted to evaluate the interpretability of the items, and items were modified as needed based on interview results. Using pilot test data, the scale demonstrated good internal consistency with a Cronbach's alpha of .86. Higher order confirmatory factor analyses supported the creation of an overall health literacy score as well as five subscale scores corresponding to the following components of health literacy: print-prose, print-document, print-quantitative, oral, and Internet (McCormack et al., 2010).

In addition to the health literacy items, we also administered the S-TOFHLA (Wallace, 2006). Baker and colleagues (1999) reported a Cronbach's alpha of .68 for numeracy and .97 for the reading comprehension items of the S-TOFHLA. The overall correlation between the S-TOFHLA and the REALM was .80.

We also asked participants to self-report their performance on the kind of skills being assessed in the survey. In particular, we asked how easy or difficult it is to remember information they read versus hear; how easy or difficult it is to understand information they read versus hear; and how easy or difficult it is to explain a health issue to their doctor, find health information they need, and locate health information on the web. Responses for each of the seven items included very difficult/difficult/ somewhat easy/very easy. Sociodemographic characteristics and selected health-related data on respondents were available from Knowledge Networks.

Statistical Methods

To identify items for the short form of the HLSI, we first reviewed the psychometric properties of the 25 items comprising the long form, including the percentages of correct responses, percentages of missing data, item response theory parameters, and factor loadings; these findings are described in McCormack and colleagues (2010). We then conducted additional analyses to test for possible differential item functioning across the following subpopulations, using an item response theory–based approach: (a) gender (male vs. female), (b) age (<60 years vs. ≥60 years), (c) race (White vs. non-White), and (d) education (more than high school vs. high school or less). The presence of differential item functioning could indicate that an item is a better discriminator of health literacy among one group than another (i.e., slope differences) or that an item is more or less difficult for one group than another (i.e., threshold differences). The differential item functioning analyses were conducted using the IRTLRDIF software program, which uses likelihood ratio tests to compare item response theory models under various parameter constraints (e.g., parameters constrained to be equal across both groups vs. slopes allowed to vary across groups) to identify slope and/or threshold-related differential item functioning (Thissen, 2001). Given the volume of statistical tests included in the differential item functioning analyses, we required a *p* value of .001 or less for significant differential item functioning to control for Type 1 error rates because of multiple comparisons.

Reviewing the results of the psychometric analyses, we selected the 10 best performing items for inclusion on the brief measure, using the following a priori criteria: (a) items should have high factor loadings and item response theory slopes, indicating good discrimination; (b) to avoid potential floor and ceiling effects, items should not have percentages correct close to 0 or 100%; (c) to ensure the measure encompasses a wide range of ability levels, the items on the scale should have a variety of item response theory thresholds and percentage of correct responses; (d) items with high rates of missing data and/or don't know responses may be confusing and/or irrelevant and will be excluded; (e) items should not demonstrate slope-related differential item functioning. In addition to the statistical results, the scale development team also reviewed item wording and selected items to ensure the content validity of the short form by including items that captured each of the five components of health literacy (print-prose, print-document, print-quantitative, oral, and Internet), as well as other critical health literacy skills while remaining within the 10-item limit. In cases where statistical and content considerations were incongruent, we consulted with experts to determine whether an item should be included. The 10-item HLSI-SF takes 5 to 10 min to administer; scores for the measure are computed as the percentage of items answered correctly.

After identifying the final set of 10 items for the short form, we conducted a one-factor confirmatory factor analysis using only the items on the short form to determine whether they grouped into one overall health literacy as expected. The factor analysis was conducted using the Mplus software program (Muthén & Muthén, 1998–2007) and incorporating the survey strata and weights to account for the complex survey design. Model fit was assessed using standard fit indices, including the comparative fit index, Tucker-Lewis fit index, and the standardized root mean square residual, with CFI of 0.95 or higher and SRMR of 0.08 or less indicating acceptable model fit (Hu & Bentler, 1999).

We then conducted item response theory analyses of the short form items with the Multilog software program (Scientific Software International, 2003), using a two-parameter logistic model for consistency with the calibration of the long form of the HLSI, and computed Cronbach's alpha to assess the internal consistency reliability of the short form. Construct validity was evaluated by conducting linear regression analyses to compare mean health literacy short-form scores by demographic characteristics and self-reported skills. For comparison purposes, similar analyses were also conducted with the long-form scores. Based on earlier results from the long form (McCormack et al., 2010), we hypothesized that participants with higher education levels and those who reported less difficulty with skills related to health literacy would have higher scores on the short form and that the short form would be moderately correlated with the S-TOFHLA.

Results

Participant characteristics are shown in Table 1; percentages are weighted using the survey weights. About half of the respondents were female, and they were distributed about equally across the four age categories and three education categories. About two thirds were White, 46% were married, and about half (51%) were employed. In terms of geography, 38% were in the South, 22% in the Midwest, 22% in the West, and 18% in the Northeast.

Ten items were selected for the health literacy short form, covering the following domains from the long form: print-prose (*n*=2), print-document (*n*=3), print-quantitative

Table 1. Mean health literacy scores, by participant characteristics

Characteristic	N	%	Health Literacy–Long Form			Health Literacy–Short Form		
			M (95% CI)	B (SE)	p	M (95% CI)	B (SE)	p
Gender								
Male	458	48	70 (68–73)	1.57 (2.18)	.469	67 (64–70)	0.77 (2.26)	.734
Female	431	52	69 (66–72)	REF		67 (63–70)	REF	
Age (years)								
18–29	180	22	69 (65–73)	1.54 (3.08)	.617	67 (62–72)	3.60 (3.32)	.279
30–44	205	25	71 (66–76)	3.53 (3.21)	.271	69 (64–73)	5.14 (3.10)	.098
45–59	255	27	71 (68–75)	4.11 (2.80)	.143	69 (65–72)	5.25 (2.82)	.063
60+	249	26	67 (64–70)	REF		63 (60–66)	REF	
Education								
More than high school	316	36	80 (78–83)	21.49 (2.49)	<.001	78 (75–80)	20.75 (2.63)	<.001
High school graduate	295	33	68 (65–72)	12.13 (2.21)	<.001	65 (61–69)	12.83 (2.38)	<.001
Less than high school	278	31	59 (55–62)	REF		57 (53–61)	REF	
Race								
White	664	64	74 (72–76)	REF		70 (68–73)	REF	
Black	83	13	56 (50–62)	−17.97 (3.47)	<.001	53 (46–60)	−17.37 (3.74)	<.001
Hispanic	80	17	65 (59–71)	−8.73 (3.51)	.013	64 (59–70)	−6.07 (3.42)	.076
Other	62	6	65 (57–72)	−9.06 (4.37)	.038	66 (58–73)	−4.74 (4.40)	.281
Marital status								
Married	489	46	73 (71–76)	7.34 (2.12)	<.001	71 (68–73)	6.81 (2.17)	.002
Not married	400	54	66 (63–69)	REF		64 (60–67)	REF	
Employment status								
Employed	470	51	74 (71–77)	REF		70 (67–74)	REF	
Retired	161	15	68 (64–71)	−6.00 (2.54)	.018	64 (61–66)	−6.90 (2.64)	.009
Disabled	81	11	55 (50–59)	−19.23 (4.03)	<.001	52 (47–56)	−18.86 (4.15)	<.001
Unemployed	177	23	68 (64–72)	−5.40 (2.73)	.048	69 (65–72)	−1.79 (2.64)	.499

(*Continued*)

Table 1. Continued

Characteristic	N	%	Health Literacy–Long Form			Health Literacy–Short Form		
			M (95% CI)	B (SE)	p	M (95% CI)	B (SE)	p
Geographic region								
Northeast	161	18	67 (62–72)	−3.79 (3.38)	.263	66 (61–70)	−3.43 (3.36)	.308
Midwest	206	22	70 (64–75)	−1.13 (3.44)	.742	66 (60–71)	−3.29 (3.65)	.368
South	338	38	70 (67–73)	−0.57 (2.80)	.838	67 (64–71)	−1.75 (2.87)	.542
West	184	22	71 (66–75)	REF		69 (64–73)	REF	

Note. REF = reference category.

196

($n=2$), oral ($n=2$), and Internet ($n=1$). Psychometric properties of the items are shown in Table 2. The percentage of correct responses ranged from 24% for Item 6 (percentage of saturated fat) to 90% for Item 2 (sign of stroke). Factor loadings for all items, except Item 6, were higher than 0.4. Similarly, as shown in Table 2 and evident from the item

Table 2. Psychometric properties of the Health Literacy Scale–Short Form

Item	Percentage correct	Factor loading	IRT Parameters	
			Slope	Threshold
PRINT-PROSE				
1. Which set of low density lipoprotein (LDL) and high density lipoprotein (HDL) levels is <u>best</u>?	66	0.57	1.20	−0.88
2. Which of the following is <u>not</u> a sign of a stroke?	90	0.71	1.84	−1.88
PRINT-DOCUMENT				
3. Which of the following entrance is closest to the elevator?	80	0.63	1.40	−1.36
4. In the example listed in the first row of the table, when should the medicine be taken?	59	0.56	1.15	−0.47
5. A person is cooking dinner for himself and he wants to include one serving from the meat and beans group. What should he choose?	75	0.68	1.59	−1.09
PRINT-QUANTITATIVE				
6. If a person is on a 2,500 calorie diet, what percent of the daily value of saturated fat would he get from one serving?	24	0.36	0.67	2.06
7. More men die from prostate cancer than from other causes. Based on the chart above, would you say this is true, false, or are you not sure?	80	0.64	1.62	−1.42
INTERNET				
8. Kate weighs 150 pounds. Which activity would burn the most calories?	54	0.61	1.36	−0.24
ORAL				
9. If a person was worried about his cough, what number should he press?	58	0.49	0.91	−0.43
10. What part of the body do lunge exercises work?	89	0.86	2.46	−1.77

Note. Factor loadings based on one-factor confirmatory factor analysis (CFI = 1.00, TLI = 1.00, RMSEA < 0.01).

characteristic curves in Figure 1, all items except Item 6 had item response theory slopes near or above 1.0, indicating good discrimination.

None of the items except Item 6 (percentage of saturated fat) demonstrated significant slope or threshold-related differential item functioning by gender, age, race, or education. Item 6 (percentage of saturated fat) did not demonstrate differential item functioning for gender or age. However, this item demonstrated significant threshold-related differential item functioning by education ($p < .001$), suggesting that it is more difficult for those with a high school education or less when compared with those with more than a high school education (i.e., some college or more). In addition, it demonstrated slope-related differential item functioning by race ($p < .001$), suggesting it can more effectively distinguish health literacy levels for White respondents than non-White respondents.

On average, participants answered 67% of the items on the short form correctly ($SD = 23\%$) compared with 70% on the long form ($SD = 22\%$). The scale demonstrated acceptable internal consistency reliability for use in making group-level comparisons with Cronbach's alpha of .70 and had a small to moderate correlation with the S-TOFHLA ($r = 0.36$). Comparisons of health literacy short- and long-form scores by participant characteristics are also shown in Table 1. Consistent across both measures, higher health literacy scores were found among those who had higher education and were married, and lower scores were found among those who were Black (vs. White) and retired or disabled (vs. employed). Those who were Hispanic or other race or unemployed had significantly lower scores on the long form, but not the short form.

As shown in Table 3, participants with poorer self-reported abilities on a range of health literacy skills had significantly lower scores on both forms of the scale ($p < .001$). The magnitudes of difference are similar between the two scales as shown
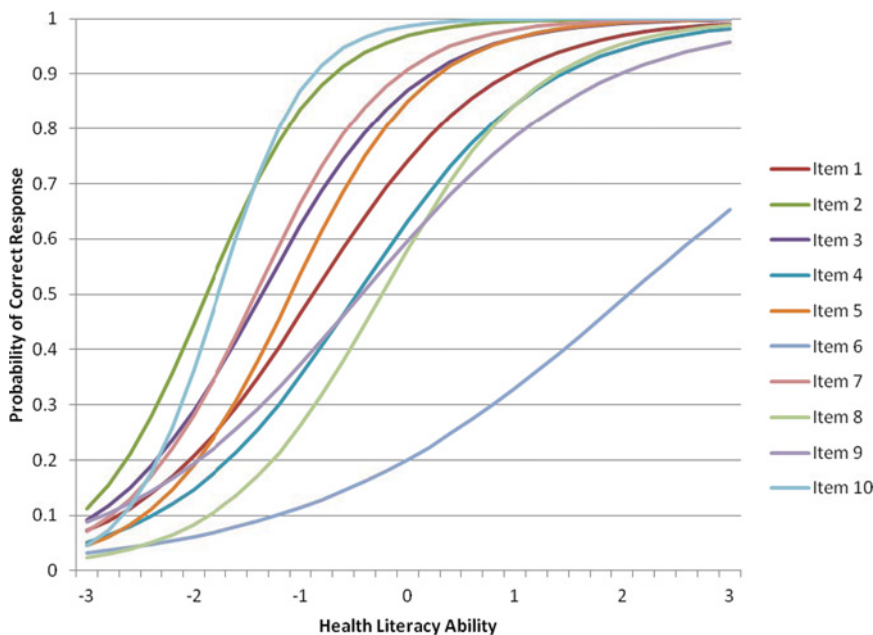

Figure 1. Item characteristic curves. (Color figure available online.)

Table 3. Mean health literacy scores, by self-reported skills

Skill	n	%	Health Literacy—Long Form			Health Literacy—Short Form		
			M (95% CI)	B (SE)	p	M (95% CI)	B (SE)	p
I am good at math								
Agree	429	46	77 (74–79)	12.59 (2.07)	<.001	74 (71–77)	12.35 (2.16)	<.001
Disagree	450	54	64 (61–67)	REF		62 (58–65)	REF	
Remembering information I read								
Easy	565	65	73 (71–76)	9.93 (2.28)	<.001	71 (68–73)	9.50 (2.32)	<.001
Difficult	317	35	63 (60–67)	REF		61 (58–64)	REF	
Remembering information I hear								
Easy	584	66	72 (70–75)	8.00 (2.42)	.001	70 (67–72)	7.48 (2.40)	.002
Difficult	295	34	64 (61–68)	REF		62 (59–66)	REF	
Understanding information I read								
Easy	665	73	75 (73–77)	18.42 (2.59)	<.001	72 (70–74)	17.08 (2.69)	<.001
Difficult	211	27	57 (52–61)	REF		55 (51–59)	REF	
Understanding information I hear								
Easy	691	77	73 (71–75)	14.89 (2.87)	<.001	71 (69–73)	15.22 (2.94)	<.001
Difficult	189	23	58 (54–63)	REF		56 (51–60)	REF	
Explaining a health issue to my doctor								
Easy	682	74	72 (70–74)	8.87 (2.77)	.001	70 (67–72)	9.62 (2.85)	<.001
Difficult	196	26	63 (59–68)	REF		60 (56–65)	REF	
Locating health information on the Internet								
Easy	654	72	74 (71–76)	13.01 (2.57)	<.001	71 (69–73)	13.33 (2.73)	<.001
Difficult	222	28	61 (56–65)	REF		58 (53–62)	REF	
Finding health information I need								
Easy	645	72	73 (71–75)	11.53 (2.67)	<.001	71 (69–73)	11.78 (2.78)	<.001
Difficult	231	28	62 (57–67)	REF		59 (54–64)	REF	

Note. REF = reference category.

199

by the regression coefficients (B). These skills encompass each of the domains covered by the scale, including print-prose (remembering and understanding information I read), print-document (finding health information I need), print-quantitative (good at math), Internet (locating health information on the Internet), and oral (remembering and understanding information I hear, explaining a health issue to a doctor). The strong relation between the lower self-reported skills and the scores on the short form of the instrument support its construct validity.

Discussion

Improving health literacy has the potential to promote more informed decision making, reduce health risks, increase prevention and wellness, and improve navigation of the health system, patient safety, patient care, and quality of life (The Calgary Charter on Health Literacy; Centre for Literacy, 2012). To track changes in health literacy over time, we need to measure and monitor it. Developing shorter versions of health literacy instruments increases the likelihood that they will be used in intervention research and other assessments, including surveillance at the local, state and national levels. Reducing the length of the instrument achieves several efficiencies related to data collection and decreases respondent burden.

We found a high correlation (.90) between the longer (25-item) and the shorter (10-item) versions of the HLSI. The average number of items answered correctly was similar between the two versions as well (70% for the HLSI vs. 67% for the HLSI-SF). The National Assessment of Adult Literacy aimed for an average of 67% correct across its items. The HLSI-SF also demonstrated acceptable internal consistency reliability and had a small to moderate correlation of the TOFHLA as expected. The HLSI-SF's correlation with the TOFLHA is slightly lower than the longer HLSI, perhaps because of the relatively higher number of items related to oral and internet skills in the short form relative to the longer form, neither of which are addressed in the TOFHLA.

We decided to retain one item (Item 6—the saturated fat item) even though it had a lower slope and factor loading and demonstrated slope-related differential item functioning by race and education because we felt that it reflected an important health literacy skill. The lower loading could be due to its quantitative emphasis. Small inconsistencies in detection of demographic differences between long and short forms may be due to elimination of items with differential item functioning from the short form.

The HSLI-SF offers many of the same advantages of the longer version with the additional benefit of taking less time to administer. Both versions are based on a conceptual framework for health literacy published in this special issue (Squiers et al., 2012), treat health literacy as a latent construct, and take a skills-based approach to measurement. They go beyond assessing what individuals can read to how people use information to manage their health and health care. The data for 25-item and 10-item versions are based on a sampling frame that covers 99% of the United States and includes individuals with higher and lower socio-economic status. As with longer form, the items in the short form reflect a variety of health related content including health promotion and disease prevention (e.g., cholesterol testing and nutrition items), treatment (e.g., medication adherence), health system navigation (e.g., hospital map).

A recent systematic review of the literature concluded that low health literacy can play an important role in the interrelation among patient characteristics, health care service use, and resulting health outcomes (Berkman et al., 2011). Using the short or

long form of the HLSI to conduct ongoing surveillance of health literacy skills could help track changes in health literacy over time and contribute to our understanding of the relation between health literacy and other factors.

References

Arozullah, A. M., Yarnold, P. R., Bennett, C. L., Soltysik, R. C., Wolf, M. S., & Ferreira, R. M., … Davis, T. (2007). Development and validation of a short-form, Rapid Estimate of Adult Literacy in Medicine. *Medical Care*, *45*, 1026–1033.

Baker, D. W., Williams, M. V., Parker, R. M., Gazmararian, J. A., & Nurss, J. (1999). Development of a brief test to measure functional health literacy. *Patient Education and Counseling*, *38*, 33–42.

Bass, P. F., III, Wilson, J. F., & Griffith, C. H. (2003). A shortened instrument for literacy screening. *Journal of General Internal Medicine*, *18*, 1036–1038.

Berkman, N. D., Davis, T. C., & McCormack, L. A. (2010). Health literacy: What is it? *Journal of Health Communication*, *15*, 9–19.

Berkman, N. D., Sheridan, S. L., Donahue, K. E., Halpern, D. J., Viera, A., Crotty, K., … Viswanath, M. (2011). *Health literacy interventions and outcomes: An updated systematic review.* Evidence Report/Technology Assessment No. 199. Retrieved from http://www.ahrq.gov/downloads/pub/evidence/pdf/literacy/literacyup.pdf

Centre for Literacy. (2012). *The Calgary Charter on Health Literacy: Rationale and core principles for the development of health literacy curricula.* Retrieved from http://www.centreforliteracy.qc.ca/sites/default/files/CFL_Calgary_Charter_2011.pdf

Davis, T. C., Long, S. W., Jackson, R. H., Mayeaux, E. J., George, R. B., Murphy, P. W., … Crouch, M. A. (1993). Rapid Estimate of Adult Literacy in Medicine: A shortened screening instrument. *Family Medicine*, *25*, 391–395.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.

Kutner, M., Greenberg, E., Jin, Y., & Paulsen, C. (2006). The health literacy of America's adults: Results from the 2003 National Assessment of Adult Literacy (NCES 2006–483) Retrieved from http://nces.ed.gov/pubs2006/2006483.pdf

McCormack, L., Bann, C., Squiers, L., Berkman, N. D., Squire, C., Schillinger, D., Hibbard, J., … (2010). Measuring health literacy: A pilot study of a new skills-based instrument. *Journal of Health Communication*, *15*(Suppl. 2), 51–71.

Muthén, L. K., & Muthén, B. O. (1998–2007). *Mplus user's guide* (5th ed.). Los Angeles, CA: Author.

Parker, R. M., Baker, D. W., Williams, M. V., & Nurss, J. R. (1995). The test of functional health literacy in adults: A new instrument for measuring patients' literacy skills. *Journal of General Internal Medicine*, *10*, 537–541.

Ratzan, S. C., & Parker, R. M. (2000). Introduction. In C. R. Seldon, M. Xorn, S. C. Ratzan, & R. M. Parker (Eds.), *National Library of Medicine current bibliographies in medicine: Health literacy (NLM Pub. No. CMB 2000-1)*. Bethesda, MD: National Institutes of Health, U.S. Department of Health and Human Services.

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine*, *127*, 966–972.

Scientific Software International. (2003). *IRT from SSI: BILOG-MG, MULTILOG, PARSCALE, TESTFACT*. Lincolnwood, IL: Author.

Squiers, L., Peinado, S., Boudewyns, V., Berkman, N., Bann, C., & McCormack, L. (2012). The health literacy skills conceptual model. Manuscript submitted for publication.

Thissen, D. (2001). IRTLRDIF v.2.0b: Software for the computation of the statistics involved in item response theory likelihood-ratio tests for differential item functioning. Retrieved from http://www.unc.edu/~dthissen/dl.html

U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion. (2010). *National Action Plan to Improve Health Literacy*. Washington, DC: Author.

Wallace, L. (2006). Patients' health literacy skills: The missing demographic variable in primary care research. *Annals of Family Medicine*, *4*, 85–86.

Weiss, B. (2009). NAAL data: To use or not to use. *Measures of health literacy: Workshop summary*. Washington, DC: The National Academies Press.

Weiss, B., Mays, M. Z., Martz, W., Castro, K. M., DeWalt, D. A., Pignone, M. P., … Hale, F. A. (2005). Quick assessment of literacy in primary care: The Newest Vital Sign. *Annals of Family Medicine*, *3*, 514–522.