University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Sociology Department, Faculty Publications

Sociology, Department of

9-2011

The self-assessed literacy index: Reliability and validity

Kristen Olson
University of Nebraska-Lincoln, kolson5@unl.edu

Jolene D. Smyth
University of Nebraska-Lincoln, jsmyth2@unl.edu

Ying Wang
University of Nebraska-Lincoln

Jennie E. Pearson
University of Nebraska-Lincoln

Follow this and additional works at: https://digitalcommons.unl.edu/sociologyfacpub

Part of the Sociology Commons

The self-assessed literacy index:
Reliability and validity

Kristen Olson,[1,2] Jolene D. Smyth,[1,2] Ying Wang,[1] and Jennie E. Pearson [1]

1. Survey Research and Methodology Program, University of Nebraska–Lincoln,
201 N. 13th St., Lincoln, NE 68588-0241, USA

2. Department of Sociology, University of Nebraska–Lincoln,
724 Oldfather Hall, Lincoln, NE 68588-0324, USA

Corresponding author — K. Olson, 703 Oldfather Hall, University of Nebraska–Lincoln,
Lincoln, NE 68588-0324, USA; fax 402 472-7764, email kolson5@unl.edu

Abstract
Literacy is associated with many outcomes of research interest as well as with respondents' ability to even participate in surveys, yet very few surveys attempt to measure it because doing so is often complex, requiring extensive tests. The central goal of this paper is to develop a parsimonious measure of respondents' reading ability that does not require a complex literacy test. We use data from the 2003 National Assessment of Adult Literacy to identify correlates of reading ability to form a literacy index. These correlates include self-assessments of one's ability to understand, read and write English, and literacy practices at home. Our literacy index reliably discerns literacy test scores above educational attainment, and the index shows high internal consistency (coefficient alpha = 0.78) and validity. The paper concludes with implications of these findings for survey research practitioners and suggestions for future research.

Keywords: literacy index, reading, survey design, data quality, scale validation

1. Introduction

Changes in survey data collection have heightened the need to understand the relationship between respondent literacy and survey data quality. One such change is that telephone surveys are losing effectiveness due to reduced coverage and response rates (Lavrakas, 2010; Tucker and Lepkowski, 2008). Over one quarter of US households are not covered by traditional landline RDD telephone surveys because they are cell-phone-only households or because they do not have telephone service (Blumberg and Luke, 2010). These changes have lead some survey researchers to transition their telephone surveys to self-administered mail surveys, replacing a random digit dial frame with an address-based frame (Link et al., 2008). This heavier reliance on self-administered mail surveys, as well as growing use of web surveys, means that the burden of reading the questionnaire and entering responses is shifted from interviewers to respondents. As such, understanding how individuals with varying literacy levels process survey questionnaires is increasingly important. The unavailability of a literacy assessment that can be included in most surveys, especially in self-administered surveys, precludes such understanding.

In 2003, the National Assessment of Adult Literacy (NAAL) estimated about 30 million American adults, or about 14% of the population, had prose literacy skills below a basic level (Kutner et al., 2007). Of these, 11 million adults, or 5% of the population, were fully nonliterate in English, leaving 9% of American adults classified as literate but with extremely low levels of literacy. These adults are able to carry out only the most basic textual search tasks, but are not necessarily able to understand the text that they read. In addition, approximately 30% of adults have only basic levels

of literacy, meaning they are able to read and write simple passages. Problems with literacy, defined as "an individual's ability to read, write, speak in English, compute and solve problems at levels of proficiency necessary to function on the job, in the family of the individual and in society" (Workforce Investment Act of 1998, Public Law 105–220 – August 7, 1998), are thus experienced by about 45% of adults in the United States.

Literacy is associated with many important economic, health and social outcomes, including educational attainment (Kirsch et al., 1993; Mellard et al., 2007), social isolation (Adkins and Ozanne, 2005), and numerous health behaviors and health status (see DeWalt et al., 2004 for systematic review; Lee, 1999; Parker, 2000). People with higher levels of literacy are more likely to be employed, work more hours, be in higher or supervisory positions and have higher incomes (Kirsch et al., 1992, 1993). Literacy is also associated with depression (Bennett et al., 2007), cognitive ability (Manly et al., 2004), and executive function (Johnson et al., 2006). In fact, literacy has been found to explain more variation in these outcomes than age or years of education (e.g., Manly et al., 2004).

Literacy is also related to nonresponse errors and measurement errors in sample surveys. Sampled persons who have difficulty reading will be unlikely or unable to participate in self-administered surveys (Groves, 2006; Dillman et al., 2009), thus increasing the risk of nonresponse error on estimates related to literacy, such as measures of health, cognitive ability, and educational attainment. Interestingly, empirical studies that measure the association between literacy and survey participation are limited, but the evidence suggests that respondents to face to face surveys (measured as health literacy; Griffin et al., 2010) and in mail surveys (measured as frequency of reading books; Hüfken, 2010) have higher literacy levels than nonrespondents. Furthermore, people with lower literacy levels provide higher rates of incorrect or inconsistent responses and fail to follow skip patterns in self-administered questionnaires more often than those with higher levels of literacy (Al-Tayyib et al., 2002; Iversen et al., 1999; Lessler and Holt, 1987], p. 262). Additionally, respondents with lower literacy levels have higher item nonresponse rates in self-administered questionnaires (Kupek, 1998; but see Williams and Swanson, 2001). Although tools such as audio-computer assisted self-interviewing (A-CASI) are available to help overcome literacy issues in self-administered components of interviewer-administered surveys (Couper and Rowe, 1996; O'Reilly et al., 1994; Couper et al., 2009), self-administered mail and web surveys do not usually have this assistance (although web surveys have the capability to add audio assistance).

Despite its association with these important outcomes, few large scale studies actually measure literacy. No single "gold-standard" literacy assessment exists, and those that do exist require interviewer-administered tests that may be too costly or too intimidating to administer. General literacy assessments measure at least one of three types of literacy skills: prose (reading and comprehending continuous texts like books and newspapers), document (processing non-continuous texts like job applications, surveys, drug or food labels) and quantitative (performing basic calculations like balancing a checkbook) (White and Dillow, 2005). Commonly used general purpose reading assessments include the Wide Ranging Achievement Test (WRAT-3), the Slosson Oral Reading Test-Revised (SORT-R) and the Peabody Individual Achievement Test (PIAT). Literacy has been assessed in health studies using the Test of Functional Health Literacy in Adults (TOFHLA) (Parker et al., 1995), the Rapid Estimate of Adult Literacy in Medicine (REALM) (Lee, 1999), or the Medical Terminology Achievement Reading Test (MART) (Hanson-Divers, 1997). Scores from these assessments measure literacy on a continuum, ranging from complete inability to read and write to high levels of proficiency (Fanta-Vagenshtein, 2008; Murray, 1995). These assessments vary in length from about two to three minutes (REALM) to up to 22 min (TOFHLA), test for general or health-specific word recognition and pronunciation, ask respondents to complete simple tasks such as reading prescription labels, and have to be administered by trained interviewers (Davis et al., 1991, 1993; Hanson-Divers, 1997; Parker et al., 1995; Lee, 1999; Manly et al., 2004; Johnson et al., 2006; DeWalt et al., 2004), making them largely infeasible for most survey research settings such as those conducted by telephone, mail or web. Although the content of the literacy assessments vary, they are highly correlated (above $r = 0.70$) (Davis et al., 1993; Baldi et al., 2009).

In lieu of a literacy assessment, proxy measures such as education are used in most surveys (Couper and Rowe, 1996; Murray, 1995; Wells et al., 1994). While education level is correlated with literacy levels (Manly et al., 2004), measures such as years of education overestimate literacy skill level by three to five reading levels (Weiss et al., 2005; Lee, 1999; Wilson, 1995). In fact, only 5% of persons with a high school degree, 31% of persons with a Bachelor's degree and 36% of persons with a graduate degree scored at the highest proficiency levels in the 2003 National Assessment of Adult Literacy (Kutner et al., 2007, p. 39).

Other proxies include self-reported literacy, presence of reading materials, reading proficiency, and reading frequency (Cooper, 1997; Murray, 1995; Rodriguez and Hagan, 1991), but are rarely assessed in surveys. One potential reason for the omission of these questions is the risk for socially desirable responses on individual items. The ability to speak, read and write English are skills considered normative for most people living in the United States (Kirsch et al., 1993; Kutner et al., 2007), especially native born US citizens. Additionally, the presence of reading material such as books and magazines in the home is a reflection of socioeconomic status (Kirsch et al., 1993) and ability to read English. As such, any one of these measures used alone is a poor indicator of literacy.

Given the prohibitive time, expense, and logistical challenges of administering an actual literacy assessment, a reliable and valid literacy scale or index using simple self-reported measures is needed. This paper develops a new measure of literacy using behavioral questions that can be implemented across multiple modes with or without an interviewer present. We develop this literacy index using the National Assessment of Adult Literacy, a survey containing objective reading assessments and subjective questions about literacy practices and behaviors. We then show the

index to have suitable face validity, construct validity, criterion validity, and predictive validity.

2. Data and methods

2.1. Data

The data for this study come from the 2003 National Assessment of Adult Literacy survey (NAAL) public use data file. The NAAL was conducted by Westat for the National Center for Education Statistics. An area probability sample of 18,186 adults ages 16 and older in the United States were interviewed in person using both English and Spanish-language questionnaires. The 2003 survey included a household sample and a prison sample; we examine only the household sample due to differences in the questionnaire for these two groups. The overall weighted household sample response rate was 60.1% (Greenberg et al., 2007). The overall weighted response rate is composed of response rates for the screener, background questionnaire, and literacy assessments (Baldi et al., 2009, p. 11–14). The technical report does not state which AAPOR response rate was used; the closest appears to be AAPOR Response Rate 6, weighted by probabilities of selection.

The NAAL measures US adults' ability to read and interpret printed and written materials using a background questionnaire and three interviewer-administered literacy assessments. The background questionnaire asks about language background, previous education, political and social participation, labor force participation, literacy practices, job training and skills, family literacy, household income and welfare participation, health, and demographics. The literacy assessments measure prose, document and quantitative literacy (White and Dillow, 2005) on a continuum from 0 to 500 with results divided into four categories: below basic, basic, intermediate and proficient.

We will focus on prose and document literacy as these skills are needed for reading survey questionnaires more than quantitative literacy. The distributions for the four literacy categories are presented in Table 1 below. About 13% of all adults and about 10% of English-speaking adults fall into the below basic category for both prose and document literacy. These respondents have the most difficulty reading basic prose or forms. Basic scores are achieved for roughly one-quarter of adults for prose and one-fifth of all adults for document literacy. Intermediate scores are obtained for about 45% of adults for prose literacy and over half of all adults in document literacy. Only about 13% of the population performed at a proficient level in each type of literacy in the NAAL (calculations and White and Dillow, 2005).

Due to the time needed to assess literacy in a population-based survey setting, the NAAL allows for estimates of literacy levels across the population, but individual literacy scores were not developed for each household respondent (Greenberg et al., 2007, p. D-2,). Instead, marginal maximum likelihood estimation (MML) accounts both for uncertainty in individual scores and for the complex survey sample design. We used the AM statistical software to conduct all MML regression estimation. This software was developed and distributed specifically for estimation of the NAAL data by the American Institutes of Research (AM Statistical Software, 2009). Marginal maximum likelihood estimation methods permit inference about the distribution of literacy in the population (that is, the margin) without directly measuring literacy for each respondent. Although this increases the amount of random measurement error in the test scores, use of MML estimation methods substantially reduces the number of questions asked of each respondent. Iterative algorithms are used to solve for the parameters of interest, such as mean literacy levels or regression coefficients in models predicting literacy scores (see Baldi et al., 2009, p. 14-8–14-15 for discussion). Additionally, we assigned multiple "plausible values" to individual respondents, similar to that of a multiple imputation, to look at individual level predictors when an appropriate MML analysis is not available. All of the analyses are weighted and account for the complex survey design.

2.2. Methods

Following traditional methods for scale development (DeVellis, 1991), we used multiple reliability and validity assessments to develop the self-assessed literacy index. These included examining construct validity, criterion validity, scale reliability, concurrent validity, and predictive validity of the index items.

First, all items in the NAAL background questionnaire were reviewed as candidate questions for the self-assessed lit-

Table 1. Distribution for four prose and document literacy categories, all adults and English-speaking adults, National Assessment of Adult Literacy, 2003. (Source: White and Dillow, 2005 and authors' calculations)

	Below basic		Basic		Intermediate		Proficient	
	%	SE	%	SE	%	SE	%	SE
Prose								
Range								
% all adults	14	0.6	29	0.6	44	0.7	13	0.5
% English-speakers	10	0.5	28	0.6	48	0.7	14	0.6
Document								
Range								
% all adults	12	0.5	22	0.5	53	0.7	13	0.6
% English-speakers	10	0.5	22	0.5	56	0.7	13	0.7

Rows may not sum to 100% due to rounding.

eracy index. Survey items were selected for further analysis if they fit two criteria: (1) the item measured general literacy, not health or other domain-specific literacy and (2) the question was asked of all household respondents. The first criterion was used to establish construct validity; the second ensured that data would be available for the next steps of index development. All variables that met these two criteria were selected for further analysis. Because questions were asked using different response scales, all of the items were dichotomized to indicate "higher" and "lower" literacy (see Table 1).

Internal reliability occurs when the items in an index are strongly correlated. After paring down the items based on the above analyses, the internal reliability of the remaining items was assessed using Kuder-Richardson 20 (KR-20) coefficients – the equivalent of Cronbach's alpha for dichotomous items. KR-20 ranges from zero to one. Although there are no clear established cutpoints for acceptable levels of KR-20 or Cronbach's alpha, values above 0.65 are considered minimally acceptable, with values above 0.70 considered acceptable (DeVellis, 1991, p. 85).

Criterion validity occurs when a scale or index is highly predictive of a gold standard measurement (Groves, 1989). Criterion validity was assessed by regressing each of the selected items on the objective literacy measures (i.e., interviewer-administered literacy assessments) in the NAAL using a marginal maximum likelihood (MML) regression. An item was considered to have criterion validity if the direction of the coefficient was consistent across models predicting document, prose, and quantitative literacy (not presented here) in univariate and multivariate analyses and the statistical significance level did not drop below 0.05 in any of the models. Given the large sample size, we required the predictor to be statistically significant across all three models. If an item was significantly associated in bivariate analyses, but not multivariate analyses at the 0.05 level, it was dropped from consideration for the index. All of the analyses were conducted separately for all households and for English-speaking households alone. We report three measures of criterion validity. First, the R^2 statistic from the MML regression predicting the raw literacy scores told us how much variance of the literacy scores are explained by the index. This is the strictest measure of criterion validity. We next treated the index as a latent continuous trait, imperfectly measured through the categorical ordinal variable (Mislevy, 1985, 1991). We examined the correlation between the latent literacy trait as measured by the index and the raw prose and document literacy scores. This analysis is conducted using the MML Table (Ordinal) procedure in the AM software. Finally although the index may not explain large amounts of variance in the raw scores, it may discriminate well across the literacy categories. Therefore, we estimated gamma coefficients to examine the association between the index and the four categories of "below basic," "basic," "intermediate" and "proficient" literacy. Gamma is a "proportional reduction in error" statistic, with "an interpretation analogous to that of r^2" (Costner, 1965, p. 350). Gamma is also a symmetric statistic, meaning that the ability to predict the SALI scores using the assessed literacy categories is improved by the same amount as the ability to predict the assessed literacy categories using the SALI. Each of these measures of criterion validity tells a slightly different story, explaining the raw scores directly, through latent variable models, and explaining the categories of literacy.

Concurrent validity "is the correlation between a measure and some criterion measured at the same point in time" (Groves, 1989, p. 23). Since education is often used as an alternative proxy for literacy, the index should be correlated with education to demonstrate concurrent validity. We also examined the association of the index with primary language (English vs. Spanish) and with age. To evaluate whether the index explains unique variance in literacy scores, we also reran our criterion validity models including education, primary language, and age as control variables.

Finally, we reevaluated whether the relationship between the index variable and outcomes of interest were of similar direction as those between literacy and outcomes of interest. We focus on ratings of self-reported health, volunteerism and employment status as examples of outcomes of interest to many literacy and social science researchers.

3. Results

3.1. Identification of possible index items

After a thorough review of the NAAL questionnaire, nine questions (Table 2) were identified that were asked of every respondent and have a probable or theoretical link to reading ability, excluding educational attainment, physical or learning disabilities, or health-specific literacy practices. These questions fell into one of two domains (1) self-assessed literacy and (2) general literacy practices.

We recoded the selected items into dichotomous variables such that a low score (0) corresponds to lower ability and a high score (1) indicates higher ability. In particular, responses to the self-rated literacy questions were coded such that responses of understanding, reading or writing English "very well" indicated higher ability and responses of "well," "not well," and "not at all" indicated lower ability. Overall, 83.2% of adults reported understanding spoken English very well, 75.6% reported reading it very well, and 70.1% reported writing English very well (Table 2). Persons who report having at least 25 books in their home or having a variety of magazines or other reading materials were coded as having higher ability. Almost 90% of adults reported having either of these reading materials. People who read newspapers or magazines, books, or letters and notes every day also were coded as having higher literacy.

3.2. Inter-item correlations and scale reliability

Next we examined the Spearman correlations among the dichotomized items to evaluate construct validity. We

Table 2. Candidate questions for the literacy index.

Q#	Question	Domain	Category	Percent	Std. error (%)
A15	With regard to English language, how well do you…				
	Understand it when it is spoken to you?	Self-assessed literacy	Very well	83.2	0.69
	Speak it?	Self-assessed literacy	Very well	76.8	0.88
	Read it?	Self-assessed literacy	Very well	75.6	0.78
	Write it?	Self-assessed literacy	Very well	70.1	0.93
E2	How often do you read … in English?				
	Newspapers or magazines	Literacy practices	Every day	48.4	0.66
	Books	Literacy practices	Every day	31.6	0.56
	Letters and notes	Literacy practices	Every day	50.6	0.79
H13	There are 25 or more books in home now	Literacy practices	True	88.2	0.49
	There is a variety of magazines and other reading materials	Literacy practices	True	89.8	0.42

The category represented here is the "higher" literacy category. The "lower" literacy categories are "well," "not well," and "not at all" for the understand, speak, read and write English questions (A15), reading newspapers or magazines, books, and letters and notes less than "every day" (E2), and answering "false" to the H13 series of questions.

expected positive correlations among the items. We also expected that correlations among items measured using similar response formats (e.g., true/false; very well, well, not well, not at all) would be stronger than correlations among items measured with different response formats due to methods effects (Alwin, 2007). Table 3 presents the correlations among the items for all adults below the diagonal in bold font and for the English-only sample above the diagonal in italics.

The items with the highest correlation were self-rated ability to read English and self-rated ability to write English (all households, $r = 0.79$, $p < .0001$; English-speaking households, $r = 0.75$, $p < .0001$). The two reading practice items (more than 25 books in the home and variety of magazines in the home) had moderate association overall ($r = 0.40$, $p < .0001$) and for English-speaking adults ($r = 0.33$, $p < 0.0001$). The correlation of $r = 0.15$ ($p < 0.0001$) between self-rated ability to speak English and reading newspapers and magazines was the weakest observed among all households. Among English-speaking adults, the weakest association was $r = 0.06$ ($p < .0001$) between reading newspapers and magazines and self-rated ability to understand spoken English.

These correlations indicate great variability in the strength of the associations across the items. To identify a more parsimonious set of items, we first calculated the Kuder-Richardson-20 coefficient (KR-20; the equivalent of Cronbach's alpha for dichotomous variables) to quantify the amount of shared variance among all of the items. Including all of the possible items yielded a KR-20 of 0.77 for all adults and 0.68 for English-speaking adults. Then, we deleted each item and examined its correlation with the total of the remaining items and the associated effect on the KR-20 coefficient. The goal of this analysis was to identify items that could be eliminated from the index because (1) their correlation with the remaining items was low and, as a result, (2) their omission from the index improved the KR-20 value among the remaining items. As can be seen in the left side of Table 4, reading newspapers and magazines had the lowest item-total correlation. For this item, the correlation with the total of all nine items for all households was 0.27

Table 3. Spearman correlation matrix for five self-reported literacy measures, all adults below the diagonal in bold font, correlation for English-only adults are above the diagonal in italic font. All adults in bold, English-speaking adults in italics.

Correlation coefficient	Understand spoken English	Speak English	Read English	Write English	Read newspapers and magazines	Read books	Read letters and notes	More than 25 books in home	Variety of magazines in home
Understand spoken English	1.00	0.64	0.58	0.52	0.06	0.11	0.10	0.10	0.08
Speak English	0.73	1.00	0.67	0.66	0.07	0.10	0.09	0.10	0.07
Read English	0.68	0.73	1.00	0.75	0.11	0.17	0.13	0.12	0.10
Write English	0.61	0.72	0.79	1.00	0.09	0.15	0.18	0.12	0.09
Read newspapers and magazines	0.17	0.15	0.19	0.16	1.00	0.15	0.18	0.10	0.13
Read books	0.19	0.17	0.23	0.21	0.19	1.00	0.19	0.16	0.10
Read letters and notes	0.19	0.16	0.20	0.19	0.21	0.22	1.00	0.13	0.12
More than 25 books in home	0.24	0.21	0.22	0.21	0.16	0.20	0.18	1.00	0.33
Variety of magazines in home	0.21	0.17	0.19	0.17	0.18	0.15	0.17	0.40	1.00

Table 4. Item-total correlation and Kuder–Richardson 20 for all households and English-speaking households for candidate SALI items.

	All candidate items				Final index			
	All adults		English-speaking adults		All adults		English-speaking adults	
	KR-20 = 0.769		KR-20 = 0.684		KR-20 = 0.779		KR-20 = 0.689	
	Corr. with total	K-R 20 if deleted	Corr. with total	K-R 20 if deleted	Corr. with total	K-R 20 if deleted	Corr. with total	K-R 20 if deleted
Understand spoken English	0.634	0.723	0.503	0.636	0.659	0.704	0.518	0.614
Speak English	0.641	0.718	0.538	0.621	–	–	–	–
Read English	0.691	0.709	0.611	0.604	0.731	0.671	0.654	0.534
Write English	0.646	0.714	0.571	0.608	0.680	0.691	0.607	0.556
Read newspapers and magazines	0.277	0.777	0.199	0.699	–	–	–	–
Read books	0.311	0.772	0.257	0.685	–	–	–	–
Read letters and notes	0.309	0.772	0.243	0.688	–	–	–	–
More than 25 books in home	0.353	0.760	0.245	0.677	0.372	0.793	0.246	0.714
Variety of magazines in home	0.317	0.764	0.215	0.681	0.344	0.797	0.227	0.716

n = 18,069 adults overall, n = 16,199 English-speaking adults.

and dropped to 0.19 for English-speaking households. Thus, this item was a candidate for removal from the index.

3.3. Criterion validity of the items

To test the criterion validity of each of the items, we examine how well each predicted assessed literacy. We started by including all nine self-assessed literacy and literacy practices items in marginal maximum likelihood regressions predicting document and prose literacy. Three of these nine items – self-rated English speaking ability, frequency of reading newspapers or magazines in English, and frequency of reading books in English – did not meet conventional levels of significance (p < .05) in at least one of the multivariate models and were thus excluded from the index (analyses not shown, but available from the authors upon request). Frequency of reading letters and notes in English was consistently significant and positive for all three literacy tests. However, additional analyses indicated that excluding this variable meaningfully improved the reliability of the index (an increase from KR-20 of 0.75 to 0.78 for all adults; increase from 0.646 to 0.69 for English-speaking adults; the right half of Table 4). Although these items were not specifically developed to be a scale, the reliability of the index is adequate using traditional cutoffs of 0.65 and 0.70 (DeVellis, 1991, p. 85).

The final index contained five variables (understand spoken English, read English, write English, more than 25 books at home, and variety of magazines at home) which were added together to form the self-assessed literacy index (SALI). The SALI ranges from 0 to 5, with a mean of 4.07 (SD = 1.87) for all adults and a mean of 4.27 (SD = 1.32) for English-speaking adults. All five variables were strong and highly significant (p < 0.0001) predictors of both prose and document literacy (results not shown; available from authors on request). This indicated that people who understand spoken English, read English, write English, and have 25 or more books or a variety of magazines and reading materials at home have significantly higher literacy scores than those who do not have these characteristics. In addition, when the five items were summed to create an index, the index was significantly and positively associated with each of the types of literacy as measured by the NAAL assessment (p < .0001 for all models; results available from authors on request). The single variable index explained roughly the same amount of variability in literacy scores as the five variables individually, ranging from 1.9% for English-speaking households on the document literacy test to 4.22% for all households on the prose literacy test.[3] When the index was treated instead as a latent trait, rather than one directly measured, the correlation, *rho*, and the explained variance in this latent trait model, measured as rho^2, rose, ranging from an explained variance of 15.9% for English-speaking households on the document literacy test to 34.0% for prose literacy for all households (see Table 5).

About 60% of all adults and about 64% of English-speaking adults had a value of 5 on the index, indicating that they provided a "high literacy" report on all five of the questions (Table 6). Yet, almost 16% of the population had a value of 0, 1 or 2 on the index, indicating they provided no more than two reports in the "high literacy" category across all five of the questions. Looking only at English-speaking adults, about 11% fell into the lowest three values on the index. In comparison, about 14% of the US adult population fell into the below basic category, 29% at the basic category, 44% to intermediate, and about 13% to proficient.

Table 6 also shows the mean document and prose literacy score as measured by the NAAL assessment for each level of the SALI for all adults and for English-speaking adults. As expected from the above regression results, there

---

3. *R*-squared measures for linear regression models estimated on data with complex sample designs and multiple imputations were not reported by the AM software. An estimated R^2 was calculated by the authors using the log-likelihood from the model.

Table 5. Explained variance of SALI index predicting prose and document literacy as continuous and categorical measures, all adults and English-speakers only, National Assessment of Adult Literacy, 2003.

	Prose		Document	
	All	English only	All	English only
Raw literacy scores				
MML R^2	4.22%	2.28%	3.27%	1.90%
Latent variable correlation (rho)	0.583	0.450	0.547	0.399
Latent variable rho^2	34.0%	20.3%	29.9%	15.9%
Literacy categories				
Gamma	0.577	0.479	0.515	0.425

Table 6. Percent of adults by SALI score, overall and English-speaking adults only, NAAL.

	Percent distribution				Mean prose literacy score				Mean document literacy score			
	All adults		English speaking adults		All adults		English speaking adults		All adults		English speaking adults	
SALI Score	Percent (%)	SE (%)	Percent (%)	SE (%)	Mean	SE	Mean	SE	Mean	Se	Mean	Se
0	2.5	0.3	0.7	0.1	142.3	4.7	190.6	7.1	140.3	4.1	182.1	6.2
1	3.9	0.2	2.2	0.2	189.2	3.6	214.9	3.8	194.3	3.5	213.8	4.4
2	10.5	0.4	8.5	0.5	234.7	2.5	246.7	2.8	239.9	1.9	248.7	2.3
3	10.1	0.5	10.4	0.5	251.4	2.1	252.7	2.1	249.0	1.9	249.6	1.8
4	13.1	0.4	13.9	0.5	274.0	2.2	274.6	2.2	267.9	2.0	268.1	1.9
5	59.9	1.0	64.2	1.0	293.9	1.3	294.0	1.3	286.5	1.2	286.3	1.7

was a monotonic relationship between the index scores and the assessed literacy levels – respondents who had a lower index score also scored lower in the NAAL prose and document literacy assessments. Likewise, respondents with the highest SALI scores tended to score higher in the prose and document literacy assessments. Notably, the group with a 0 on the SALI for English-speaking adults and with a 0 or 1 for all adults had mean document and prose literacy scores that fell below the "below basic" cutoff (210 for prose and 205 for document) in the NAAL assessment of both tests. That is, the lower values of the index reliably discern those who have the most difficulty with reading. Moreover, English-speaking and all adults with SALI scores of 4 or 5 had mean NAAL assessment scores that fell between the "intermediate" (265 for prose and 250 for document) and "proficient" (340 for prose and 335 for document) cutoff in both tests. Thus, the higher values of the index reliably discern those who have the least difficulty with reading.

Figure 1, Panel A, illustrates the ability of the SALI to discriminate between NAAL literacy levels. The percentage of adults in each proficiency level of prose literacy is displayed for each level of the SALI for all adults and English-speaking adults. Respondents who scored in the lowest SALI category were most likely to fall in the below basic category in the NAAL prose literacy assessment while very few respondents who were assessed at the intermediate or proficient levels fall into this category. Likewise, those who fall into the highest SALI categories were likely to score in the intermediate and proficient NAAL levels, while very few respondents scoring at the basic or below basic levels were found in these higher categories. This pattern is consistent for document literacy (not shown), for all adults, and for English-speaking adults. When the literacy categories were examined rather than the raw scores, the association between the index and the literacy proficiency categories was strong, ranging from a gamma of 0.425 for document literacy among English speaking adults to 0.577 for prose literacy among all adults. That is, we improved our predictions in the prose literacy categories by 57.7% by using the SALI over not using it at all (Costner, 1965; Kviz, 1981).

Panel B of Figure 1 illustrates the distribution of SALI scores within each assessed prose literacy category for all adults and English-speaking adults. Although only 2.5% of adults (0.7% of English-Speaking adults) had a SALI score of zero, 13% of adults in the "below basic" category had this score. In fact, 30.4% of persons in the "below basic" category had the bottom two SALI scores of 0 or 1, falling to 5.9% of those in the "basic" category, 1.0% of those in the "intermediate" category, and about 0.1% of those in the "proficient" category. Thus, those who scored at the bottom of the SALI distribution are highly unlikely to have more than "basic" prose literacy levels. Similarly, the proportion of persons in the top two SALI categories decreased as the assessed literacy levels decreased. Almost all of the adults (94.8% of all adults; 95.2% of English-speaking adults) who were assessed to be in the "proficient" literacy category had a value of 4 or 5 on the SALI, monotonically declining to 31.2% of those in the "below basic" category having these SALI scores. Those who scored at the top of the SALI distribution were found across the literacy spectrum, but were more likely to be among those with higher literacy levels.

3.4. Concurrent validity

Concurrent validity is the association between a measure and other constructs to which it theoretically should be related (Groves, 1989). To examine the concurrent validity of the SALI, we examined the correlation between the in-
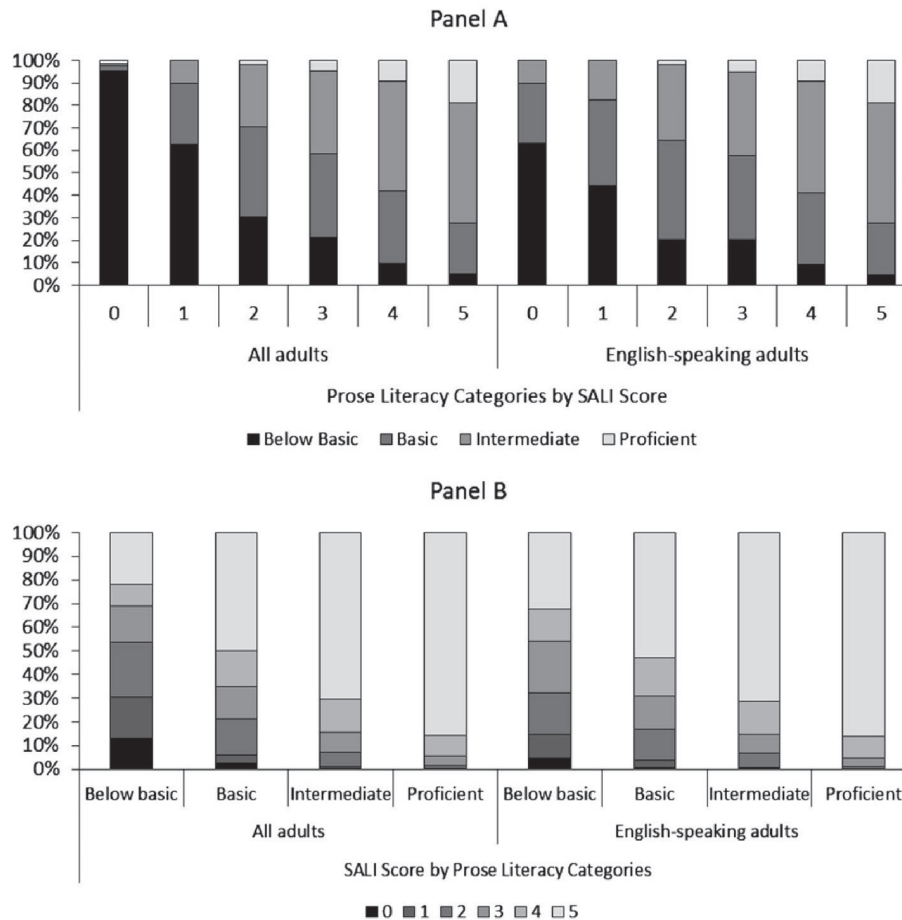
Figure 1. Percentage of adults in each category of prose literacy for each level of the index for all households and English-speaking households.

dex and education, primary language, and age. The SALI was strongly and positively associated with language and education ($r = 0.51$ and $r = 0.36$, $p < .0001$, respectively), but not significantly associated with age ($r = -0.005$, $p = 0.54$). Moreover, the association between the index and education was greater than the correlation between language and education ($r = 0.20$, $p < .0001$), or between age and either language or education ($r = 0.07$, $p < .0001$ and $r = 0.13$, $p < .0001$, respectively). Similar patterns were seen when looking only at English-speaking adults, with a significant, positive correlation between education and the SALI ($r = 0.30$, $p < .0001$) and a weak, negative correlation between age and the SALI ($r = -0.05$, $p < .0001$).

Given the aforementioned association between education and literacy that others have reported (Kirsch et al., 1993; Kutner et al., 2007)) and were found in these data, one concern that arises is that the association between the SALI and NAAL-assessed literacy scores – the test of criterion validity – is spurious. To examine this hypothesis, we reestimated the regression models predicting NAAL assessed literacy with the SALI index, controlling for age, education, and language (when looking at all households – Table 7). When looking at all adults, the magnitude of the SALI coefficient predicting document literacy decreased from beta = 20.8 (SE = 0.67) when it is the only variable in the model to beta = 11.9 (SE = 0.51) and from beta = 23.8 (SE = 0.72) to 10.2 (SE = 0.49) when predicting prose literacy, reflecting the shared variance of the index with education and language. Yet, the index remained a positive and highly significant predictor ($p < .0001$) of NAAL-assessed literacy in all models. These results indicated that despite the association between education and the SALI, the index continued to explain unique variance in literacy, above and beyond educational levels.

3.5. Predictive validity

To be useful, the SALI literacy index should predict outcomes in a direction similar to that of the assessed literacy scores. To examine this, we selected outcome variables from the three domains of health, economics, and social participation. To examine the relationship between the SALI and health, we look at self-reported health status, reported on a five point scale from poor to excellent. In the domain of social participation, we look at volunteerism. In the economic domain, we predict the probability of being unemployed or not in the labor force compared to being em-

Table 7. Coefficients and standard errors from marginal maximum likelihood regression with index in one model for all adults and English only adults.

	All adults				English-speaking adults			
	Document literacy		Prose literacy		Document literacy		Prose literacy	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Intercept	141.52	2.88	160.03	3.45	219.98	2.99	202.54	2.97
SALI index	11.90	0.51	10.16	0.49	9.17	0.51	10.94	0.53
English-language	56.84	2.50	40.32	3.45	–		–	
Age	−7.39	0.44	−5.92	0.49	−10.11	0.40	−7.34	0.46
Education	11.27	0.27	12.33	0.28	10.15	0.23	11.20	0.27
R^2	8.99%		8.99%		4.40%		6.46%	

Table 8. Linear regression coefficients and standard errors predicting self-reported health status, and logistic regression coefficients and standard errors predicting volunteerism, SALI vs. assessed literacy scores.

	All adults						English-speaking adults					
	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
Self-rated health	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
Prose Literacy	–	–	0.17	0.02	0.14	0.02	–	–	0.18	0.02	0.14	0.02
SALI index	0.11	0.01	–	–	0.09	0.01	0.13	0.01	–	–	0.10	0.01
Education	0.11	0.005	0.09	0.006	0.09	0.01	0.11	0.005	0.09	0.007	0.09	0.01
Age	−0.22	0.01	−0.20	0.007	−0.20	0.01	−0.21	0.01	−0.20	0.008	−0.20	0.01
Language	−0.20	0.05	−0.20	0.05	−0.34	0.06	–	–	–	–	–	–
Volunteerism												
Prose Literacy	–	–	0.31	0.04	0.25	0.04	–	–	0.29	0.04	0.24	0.04
SALI index	0.21	0.02	–	–	0.16	0.03	0.19	0.02	–	–	0.15	0.03
Education	0.18	0.01	0.14	0.01	0.13	0.01	0.18	0.01	0.15	0.02	0.14	0.01
Age	−0.04	0.02	−0.01	0.02	−0.01	0.02	−0.04	0.02	−0.01	0.02	−0.01	0.02
Language	0.62	0.12	0.64	0.02	0.38	0.14	–	–	–	–	–	–

Table 9. Multinomial logistic regression coefficients and standard errors predicting employment status, SALI vs. assessed literacy scores, all adults.

	Model 1				Model 2				Model 3			
	Unemployed		Out of the labor force		Unemployed		Out of the labor force		Unemployed		Out of the labor force	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
Prose literacy	–	–	–	–	−0.17	0.07	−0.14	0.03	−0.19	0.08	−0.17	0.04
SALI index	−0.003	0.047	0.018	0.020	–	–	–	–	0.039	0.054	0.054	0.026
Education	−0.15	0.02	−0.26	0.01	−0.11	0.03	−0.23	0.02	−0.11	0.03	−0.23	0.02
Age	−0.21	0.05	0.53	0.03	−0.23	0.06	0.51	0.03	−0.23	0.06	0.51	0.03
Language	0.50	0.23	0.13	0.11	0.78	0.24	0.41	0.12	0.71	0.28	0.31	0.12

ployed. We focus on the prose literacy scores in the following analyses.

Table 8 shows that NAAL-assessed prose literacy was a significant and positive predictor (beta = 0.17, se = 0.02) of self-reported health – persons who have higher literacy also report being in better health – and of being a volunteer (beta = 0.31, se = 0.04). We see an identical pattern for the SALI, with a positive and significant relationship with health status (beta = 0.11, se = 0.01) and volunteerism (beta = 0.21, se = 0.02). The two variables – although not identical – clearly share common variance, with the magnitude of the coefficients reduced when both are entered in the model.

The pattern is slightly different with employment status. Table 9 shows multinomial logistic regression coefficients predicting being unemployed or out of the labor force compared to being employed. NAAL-assessed prose literacy was strongly associated with employment status (beta = −0.17, se = 0.07 for unemployed vs. employed; −0.14, se = 0.03 for out of the labor force vs. employed). Surprisingly, the SALI was not a statistically significant predictor of being unemployed or out of the labor force (beta = −0.003, se 0.047 for unemployed vs. employed; beta = 0.018, se = 0.020 for out of the labor force vs. employed). Thus, while the SALI was a good predictor in the health and social participation domains, there may be limitations in the economic domain.

4. Discussion and conclusion

With increasing numbers of surveys being conducted in self-administered modes, it is important to understand

how different levels of literacy might affect responses. Survey researchers can work to improve the survey process for those who might struggle to answer survey questions and also evaluate whether people with low literacy are represented in their respondent pool. Those who are studying outcomes linked to literacy such as social participation, social isolation, health, and mental health will need to be particularly vigilant about ensuring that individuals from all levels of the literacy continuum make it into their respondent pool. However, since current literacy assessments require complex interviewer administration, few studies collect information about respondents' literacy skills. A simpler measure that can be included in either interviewer- or self-administered surveys is badly needed.

This paper reports efforts to develop and validate such a measure, the Self-Administered Literacy Index. Developed on a nationally-representative sample of adults in the United States, the SALI shows reasonable reliability, criterion validity, concurrent validity and predictive ability. Moreover, the SALI discriminates literacy levels among English speakers, a group frequently ignored in other examinations of literacy.

The SALI cannot completely replace other literacy assessments. Although it is associated with prose and document literacy measures in the NAAL, it is a relatively coarse measure. Researchers who need the detail that comes with more traditional assessment methods will find the SALI lacking as it does not explain extremely high amounts of variability in the raw NAAL prose and document literacy scores. It does, however, show good explanatory value for the coarser literacy categories of Below Basic, Basic, Intermediate and Proficient. In addition, the new index measures literacy on a continuum rather than as a literate vs. illiterate dichotomy. Thus, while it is a somewhat coarse measure, it does provide a viable alternative for collecting information about literacy levels in studies where great detail is not needed or when it is not feasible to have interviewers administer complex literacy assessments.

Because it uses only five simple questions, the SALI could easily be included in surveys without adding a great deal of respondent burden or requiring too much extra space. Although we do not have time estimates for how long these five questions took to be administered in the NAAL, a rough estimate of two to four questions per minute would indicate that these five questions would add between one and two minutes to an interviewer-administered survey. This added length is much less than that for adding traditional literacy assessments. Although the NAAL did not ask the SALI items consecutively, as a set of conceptually related items, the SALI questions could be added as a series of questions in any questionnaire. Alternatively, the SALI items could be separated – such as in the NAAL – into questions about leisure activities (of which reading is one type) and questions about general skills (of which English speaking, reading and writing are types of skills). Moreover, given its brevity, the SALI is a measure that can be obtained on each respondent. In the NAAL, individual literacy scores for each respondent could not be calculated given the desire to limit respondent burden. In contrast, the SALI questions were asked of all respondents.

An additional strength of the SALI is that interviewers do not need to receive special training. In the NAAL, interviewers were trained on how to conduct the literacy assessments using both the computerized instrument and paper booklets, with different procedures for those respondents with the lowest literacy levels (Baldi et al., 2009, pp. 1-9–1-10). The routine standardized interviewer training conducted at most survey organizations is sufficient for reading the SALI questions as written. This reduces the cost of interviewer training in addition to minimizing the risk of interviewer errors. Furthermore, in settings where measures such as the TOFHLA, REALM, or MART have been used, these five questions could cue health professionals to provide support for their patients or clients in understanding and following prescriptions or brochures on healthy behaviors, among others.

A potential drawback is that the SALI questions themselves may be subject to measurement error such as social desirability biases in self-reports about English-language ability. Furthermore, the reports about the presence of 25 or more books, a variety of magazines or other reading materials, and the frequency of reading newspapers, books, or letters or notes may be inaccurate. These inaccuracies may arise from social desirability, from difficulty in estimating the frequency of these behaviors, or from complex terms such as "variety." Although the use of multiple questions helps to overcome this limitation, it is possible that each question suffers from some of these errors. To the extent that the self-reports are systematically biased in the direction of "higher literacy" reports, persons with lower assessed literacy levels will have inflated SALI scores. This might explain the lack of association between the SALI and the economic domain of employment and the frequency of high SALI scores among those with "below basic" literacy levels.

Despite the imperfections, the SALI operationalizes a concept (literacy) that is generally measured by a different proxy, education. In our multivariate models, the SALI continues to explain literacy even after accounting for educational differences. This indicates that the SALI measures a component of literacy distinct from educational attainment (i.e., that is not accounted for in studies using only the education proxy). Educational attainment clearly measures literacy, but also is a measure of a host of other socioeconomic characteristics that are distinct from literacy. Thus, the SALI allows literacy to be measured in a more direct way without the burden of a full assessment.

The development of the SALI presented here needs to be replicated in other studies and survey modes. As detailed, one of the goals in developing this measure is that it could be transported to other surveys, including self-administered surveys. More research needs to be done to determine whether the SALI maintains its reliability and validity when administered in these modes. For instance, whether similar levels of reliability and validity are obtained in self-administered modes where literacy problems are most salient could be examined by including these questions in mail or web surveys. Additionally, empirical examination of reliability and validity when these questions are asked sequentially compared to broken into different sections of the questionnaire would further inform the use of this index. Finally, whether the SALI is also predictive of data quality outcomes such as item nonresponse may prove useful

for designing questionnaires for lower literacy respondents. The ability to measure literacy in a wide variety of surveys and survey modes opens up countless possibilities for understanding social phenomena that are related to literacy and for improving our methodological understanding about how literacy is related to survey data quality.

References

Adkins and Ozanne, 2005 • N. R. Adkins and J. L. Ozanne, The low literate consumer, *Journal of Consumer Research* **32** (1) (2005), pp. 93–105; doi: 10.1086/429603.

Al-Tayyib et al., 2002 • A. A. Al-Tayyib, S. M. Rogers, J. N. Gribble, M. Villarroel, and C. F. Turner, Effect of low medical literacy on health survey measurements, *American Journal of Public Health* **92** (9) (2002), pp. 1478–1480.

Alwin, 2007 • D. F. Alwin, *Margins of Error: A Study of Reliability in Survey Measurement,* John Wiley and Sons, Hoboken, NJ (2007).

AM Statistical Software, 2009 • AM Statistical Software (Beta version 0.06.00). American Institutes for Research, Washington, DC; retrieved November 18, 2009 from http://am.air.org/

Baldi et al., 2009 • S. Baldi et al. (eds.), Technical Report and Data File User's Manual for the 2003 National Assessment of Adult Literacy (NCES 2009-476). US Department of Education, National Center for Education Statistics. US Government Printing Office, Washington, DC, 2009.

Bennett et al., 2007 • I. M. Bennett, J. F. Culhane, K. F. McCollum, L. Mathew, and I. T. I. Elo, Literacy and depressive symptomatology among pregnant Latinas with limited English proficiency, *American Journal of Orthopsychiatry* **77** (2) (2007), pp. 243–248 doi: 10.1037/0002-9432.77.2.243.

Blumberg and Luke, 2010 • S. J. Blumberg and J. V. Luke, Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, July–December 2009. National Center for Health Statistics (May 12, 2010); http://0-www.cdc.gov.library.unl.edu/nchs/data/nhis/earlyrelease/wireless201005.pdf

Cooper, 1997 • J. D. Cooper, *Literacy: Helping Children Construct Meaning* (third ed.), Houghton Mifflin Company, Boston (1997).

Costner, 1965 • H. L. Costner, Criteria for measures of association, *American Sociological Review* **30** (3) (1965), pp. 341–353.

Couper and Rowe, 1996 • M. P. Couper and B. Rowe, Evaluation of a computer-assisted self-interview component in a computer-assisted personal interview survey, *Public Opinion Quarterly* **60** (1) (1996), pp. 89–106; http://poq.oxfordjournals.org/cgi/reprint/60/1/89

Couper et al., 2009 • M. P. Couper, R. Tourangeau, and T. Marvin, Taking the audio out of audio-casi, *Public Opinion Quarterly* **73** (2009), pp. 281–303.

Davis et al., 1991 • T. Davis, M. Crouch, S. Long, R. Jackson, P. Bates, R. George, and L. Bairnsfather, Rapid assessment of literacy levels of adult primary care patients, *Family Medicine Journal* **23** (6) (1991), pp. 433–435.

Davis et al., 1993 • T. C. Davis, S. W. Long, R. H. Jackson, E. J. Mayeaux, R. B. George, P. W. Murphy, and M. A. Crouch, Rapid estimate of adult literacy in medicine: A shortened screening instrument, *Family Medicine Journal* **25** (6) (1993), pp. 391–395.

DeVellis, 1991 • R. F. DeVellis, Scale development: theory and applications, *Sage Publications* **15** (4) (1991), pp. 425–426; doi: 10.1177/014662169101500413.

DeWalt et al., 2004 • D. A. DeWalt, N. D. Berkman, S. Sheridan, K. N. Lohr, and M. P. Pignone, Literacy and health outcome: A systematic review of the literature, *Journal of General Internal Medicine* **19** (12) (2004), pp. 1228–1239; doi: 10.1111/j.1525-1497.2004.40153.x.

Dillman et al., 2009 • D. A. Dillman, J. D. Smyth, and L. M. Christian, *Internet, Mail and Mixed-Mode Surveys: The Tailored Design Method* (3rd ed.), John Wiley & Sons, Hoboken, NJ, 2009.

Fanta-Vagenshtein, 2008 • Y. Fanta-Vagenshtein, How illiterate people learn: Case study of Ethiopian adults in Israel, *The Journal of Literacy and Technology* **9** (3) (2008), pp. 24–55; http://www.literacyandtechnology.org/volume9_3/jlt_v9_3_Fanta-Vagenshtein.pdf

Greenberg et al., 2007 • E. Greenberg, Y. Jin, and S. White, National Assessment of Adult Literacy, 2003: Public-use data file user. National Center for Education Statistics, 2007; http://eric.ed.gov:80/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/29/e5/50.pdf

Griffin et al., 2010 • J. M. Griffin, M. R. Partin, S. Noorbaloochi, J. P. Grill, S. Saha, and A. Snyder *et al.*, Variation in estimates of limited health literacy by assessment instruments and non-response bias, *Journal of General Internal Medicine* **25** (7) (2010), pp. 675–681.

Groves, 1989 • R. M. Groves, *Survey Errors and Survey Costs,* John Wiley and Sons, New York (1989).

Groves, 2006 • R. M. Groves, Nonresponse Rates and Nonresponse Bias in Household Surveys, *Public Opinion Quarterly* **70** (5) (2006), pp. 646–675.

Hanson-Divers, 1997 • E. C. Hanson-Divers, Developing a medical achievement reading test to evaluate patient literacy skills: A preliminary study, *Journal of Health Care for the Poor and Underserved* **8** (1) (1997), pp. 56–69.

Hüfken, 2010 • V. Hüfken, Supplementary questionnaire and nonresponse-results from the German ISSP survey, *Quality & Quantity* **44** (4) (2010), pp. 607–622.

Iversen et al., 1999 • R. R. Iversen, F. F. Furstenberg Jr., and A. A. Belzer, How much do we count? Interpretation and error-

making in the decennial census, *Demography* **36** (1999), pp. 121–134.

Johnson et al., 2006 • A. S. Johnson, L. J. Flicker, and P. A. Lichtenberg, Reading ability mediates the relationship between education and executive function tasks, *Journal of the International Neuropsychological Society* **12** (2006), p. 64; doi: 10.1017/S1355617706060073.

Kirsch et al., 1992 • I. S. Kirsch, A. Jungeblut, and A. Campbell, *Beyond the School Doors: The Literacy Needs of Job Seekers Served by the US Department of Labor.* US Department of Labor, Employment and Training Administration, 1992.

Kirsch et al., 1993 • I.S. Kirsch, A. Jungeblut, L. Jenkins and A. Kolstad, *Adult Literacy in America: A First Look at the Results of the National Adult Literacy Survey,* Office of Educational Research and Improvement, US Dept. of Education, Washington, DC, 1993.

Kupek, 1998 • E. Kupek, Determinants of Item Nonresponse in a Large National Sex Survey, *Archives of Sexual Behavior* **27** (6) (1998), pp. 581–594.

Kutner et al., 2007 • M. Kutner, E. Greenberg, Y. Jin, B. Boyle, Y. Hsu, and E. Dunleavy, *Literacy in Everyday Life: Results from the 2003 National Assessment of Adult Literacy* (NCES 2007–480), National Center for Education Statistics, US Department of Education, Washington, DC, 2007.

Kviz, 1981 • F. J. Kviz, Interpreting proportional reduction in error measures as percentage of variation explained, *The Sociological Quarterly* **22** (1981), pp. 413–420.

Lavrakas, 2010 • P. J. Lavrakas, Telephone surveys. In: J. D. Wright and P. V. Marsden, eds., *Handbook of Survey Research*, Elsevier, San Diego, 2010.

Lee, 1999 • P. P. Lee, Why literacy matters: Links between reading ability and health, *Archives of Ophthalmology* **117** (1) (1999), pp. 100–103 (Chapter 14). doi:10.1001/archopht.117.1.100.

Lessler and Holt, 1987 • J. T. Lessler and M. Holt, Using response protocols to identify problems in the U.S. Census long form, *Proceedings of the Survey Research Methods Section*, American Statistical Association, Alexandria, VA (1987), pp. 262–266.

Link et al., 2008 • M. W. Link, M. P. Battaglia, M. R. Frankel, L. Osborn, and A. H. Mokdad, A comparison of address-based sampling (ABS) versus random-digit dialing (RDD) for general population surveys, *Public Opinion Quarterly* **72** (1) (2008), pp. 6–27; doi 10.1093/poq/nfn003.

Manly et al., 2004 • J. J. Manly, D. Byrd, P. Touradji, D. Sanchez, and Y. Stern, Literacy and cognitive change among ethnically diverse elders, *International Journal of Psychology* **39** (2004), pp. 47–60.

Mellard et al., 2007 • D. Mellard, M. B. Patterson, and S. Prewett, Reading practices among adult education participants, *Reading Research Quarterly* **42** (2) (2007), pp. 188–213; doi: 10.1598/RRQ.42.2.1.

Mislevy, 1985 • R. J. Mislevy, Estimation of latent group effects, *Journal of the American Statistical Association* **80** (392) (1985), pp. 993–997.

Mislevy, 1991 • R. J. Mislevy, Randomization-based inference about latent variables from complex samples, *Psychometrika* **56** (2) (1991), pp. 177–196.

Murray, 1995 • T. S. Murray, *Proxy Measurement of Adult Basic Skills: Lessons from Canada.* National Center on Adult Literacy, Philadelphia, PA, 1995; http://www.literacy.org/products/ncal/pdf/TR9418.pdf

O'Reilly et al., 1994 • J. M. O'Reilly, M. L. Hubbard, J. T. Lessler, P. P. Biemer, and C. F. Turner, Audio and video computer assisted self-interviewing: Preliminary tests of new technologies for data collection, *Journal of Official Statistics* **10** (1994), pp. 197–214.

Parker, 2000 • R. Parker, Health literacy: A challenge for American patients and their health care providers, *Health Promotion International* **15** (4) (2000), pp. 277–283; doi: 10.1093/heapro/15.4.277.

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

Rodriguez and Hagan, 1991 • N. P. Rodriguez and J. S. Hagan, *Investigating Census Coverage and Content among the Undocumented: An Ethnographic Study of Latin Tenants in Houston, Texas.* Ethnographic Evaluation of the 1990 Decennial Census Report #3, 1991; http://www.censusbureau.biz/srd/papers/pdf/ev91-3.pdf

Tucker and Lepkowski, 2008 • C. Tucker and J. M. Lepkowski, Telephone survey methods: Adapting to change. In: J. M. Lepkowski, C. Tucker, J. M. Brick, E. D. de Leeuw, L. Japec, and P. J. Lavrakas *et al.*, eds., *Advances in Telephone Survey Methodology*, Wiley, New York, 2008, pp. 1–26.

Weiss et al., 2005 • B. D. Weiss, M. Z. Mays, W. Martz, K. M. Castro, D. A. DeWalt, M. P. Pignone, J. Mockbee, and F. A. Hale, Quick assessment of literacy in primary care: The newest vital sign, *Annals of Family Medicine* **3** (6) (2005), pp. 514–522 doi: 10.1370/afm.405.

Wells et al., 1994 • James A. Wells, Donna Ruscavage, Barbara Parker, and Lynne McArthur, Literacy of women attending family planning clinics in Virginia and reading levels of brochures on HIV prevention, *Family Planning Perspectives* **26** (1994), pp. 113–131.

White and Dillow, 2005 • S. White and S. Dillow, Key Concepts and Features of the 2003 National Assessment of Adult Literacy (NCES 2006–471), US Department of Education, National Center for Education Statistics, Washington, DC, 2005.

Williams and Swanson, 2001 • S. A. Williams and M. S. Swanson, The effect of reading ability and response formats on patients' abilities to respond to a patient satisfaction scale, *The Journal of Continuing Education in Nursing* **32** (2) (2001), pp. 60–67.

Wilson, 1995 • F. Wilson, Measuring patients' ability to read and comprehend: A first step in patient education, *Nurse Connections* **8** (1995), pp. 17–25.

Workforce Investment Act, 1998 • Workforce Investment Act of 1998; http://www.doleta.gov/regs/statutes/wialaw.txt (accessed February 24, 2009).