

University of San Diego

Digital USD

School of Nursing and Health Science: Faculty
Scholarship

School of Nursing and Health Science

2023

Comparison of Newest Vital Sign and Brief Health Literacy Screen scores in a large, urban Hispanic cohort

Adriana Arcia

University of San Diego, aarcia@san Diego.edu

Anthony T. Pho

Columbia University

Maichou Lou

Columbia University

Suzanna Bakken

Columbia University

Follow this and additional works at: https://digital.sandiego.edu/nursing_facpub



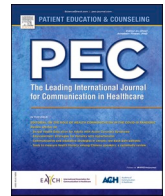
Part of the [Nursing Commons](#)

Digital USD Citation

Arcia, Adriana; Pho, Anthony T.; Lou, Maichou; and Bakken, Suzanna, "Comparison of Newest Vital Sign and Brief Health Literacy Screen scores in a large, urban Hispanic cohort" (2023). *School of Nursing and Health Science: Faculty Scholarship*. 47.

https://digital.sandiego.edu/nursing_facpub/47

This Article is brought to you for free and open access by the School of Nursing and Health Science at Digital USD. It has been accepted for inclusion in School of Nursing and Health Science: Faculty Scholarship by an authorized administrator of Digital USD. For more information, please contact digital@san Diego.edu.



Comparison of Newest Vital Sign and Brief Health Literacy Screen scores in a large, urban Hispanic cohort

Adriana Arcia^{a,*}, Anthony T. Pho^{a,3,4}, Maichou Lor^{a,5,7}, Suzanne Bakken^{a,b,6}

^a School of Nursing, Columbia University, 560 West 168th Street, MC6, New York, NY 10032, USA

^b Department of Biomedical Informatics, Columbia University, 622 W. 168th Street, PH20 3720, New York, NY 10032, USA

ARTICLE INFO

Keywords:

Health literacy
Psychometrics
Hispanic or Latino

ABSTRACT

Objective: Prior studies comparing subjective and objective health literacy measures have yielded inconsistent results. Our aim was to examine the concordance between Newest Vital Sign (NVS) and Brief Health Literacy Screen (BHLS) scores in a large cohort of English- and Spanish-speaking urban Hispanic adults.

Methods: Item means, standard deviations, corrected-item total correlations, Cronbach's alpha, and Spearman correlations and area under receiver operating characteristic (AUROC) curve analysis were used to compare NVS and BHLS items and total scores.

Results: $N = 2988$ ($n = 1259$ English; $n = 1729$ Spanish). Scores on both measures demonstrated good internal consistency (NVS: $\alpha = .843$ English, $.846$ Spanish; BHLS: $\alpha = .797$ English, $.846$ Spanish) but NVS items had high difficulty; more than half of respondents scored 0. Measures were only weakly correlated ($r_s = .21$, $p < .001$, English; $r_s = .19$, $p < .001$, Spanish). The AUROC curves were .606 (English) and .605 (Spanish) for discriminating the lowest NVS scoring category.

Conclusion: Subjective health literacy scores were poor predictors of objective scores. Objective scores demonstrated floor effects, precluding discrimination at low levels of the trait continuum.

Practice implications: Subjective health literacy scores may fail to identify individuals with limited health literacy.

1. Introduction

The Newest Vital Sign (NVS) is an objective measure of health literacy, frequently used in research because of its brevity and purported criterion-related validity in relation to longer “gold standard” measures like the Test of Functional Health Literacy in Adults (TOFHLA, and its short form S-TOFHLA) and the Rapid Estimate of Adult Literacy in

Medicine (REALM) [1]. The Brief Health Literacy Screen (BHLS) [2] is a subjective measure of health literacy that is attractive for use in clinical practice and research because it can be self-administered, is even shorter than the NVS, and there is some evidence that it yields scores that are good predictors of NVS scores [3]. However, the psychometric performance of these measures—individually and in relation to each other—has varied by population and language of administration [4,5] and thus may not be

Abbreviations: NVS, Newest Vital Sign; BHLS, Brief Health Literacy Screen; TOFHLA, Test of Functional Health Literacy in Adults; REALM, Rapid Estimate of Adult Literacy in Medicine; AUROC, Area under the receiver operating characteristic; C-ITC, Corrected item-total correlation; WICER, Washington Heights/Inwood Informatics Infrastructure for Community-Centered Comparative Effectiveness Research.

* Correspondence to: University of San Diego, 5998 Alcalá Park, BINR 321, San Diego, CA 92110, USA.

E-mail addresses: adrianaxarcia@gmail.com (A. Arcia), apho@stanford.edu (A.T. Pho), mlor2@wisc.edu (M. Lor), sbh22@columbia.edu (S. Bakken).

¹ orcid.org/0000-0002-4773-7195

² Adriana Arcia is now at the Hahn School of Nursing and Health Science, University of San Diego.

³ 0000-0002-9959-3541

⁴ Anthony T. Pho is now at the School of Medicine, Stanford University.

⁵ 0000-0001-8451-4364

⁶ 0000-0001-6202-6001

⁷ Maichou Lor is now at the School of Nursing, University of Wisconsin-Madison.

<https://doi.org/10.1016/j.pec.2023.107628>

Received 15 September 2022; Received in revised form 28 December 2022; Accepted 9 January 2023

Available online 10 January 2023

0738-3991/© 2023 Elsevier B.V. All rights reserved.

equally appropriate for use with all populations and sub-populations. Hispanic/Latinx⁸ people are the largest ethnic minority group in the US representing 18.5% of the total population but are not a monolith; there is cultural heterogeneity related to countries of origin [6]. Therefore, it is important to understand how these commonly-used health literacy measures perform within individual subgroups to ensure the validity and rigor of any attendant health literacy research.

1.1. Study purpose

The purpose of this study is to describe the psychometric characteristics of NVS and BHLS scores in English and Spanish in a large cohort of English- and Spanish-speaking urban Hispanic respondents, and to examine the concordance between them.

1.2. Newest Vital Sign (NVS)

The NVS is an objective, interviewer-administered measure of health literacy based on six questions about a nutrition label on a container of ice cream [1]. Administration takes 3–5 minutes. The NVS has an emphasis on numeracy, as the first four items require simple numerical calculations (e.g., multiplying the calories per serving by the total number of servings). Item 5 asks if it would be safe for a person with a list of four allergies to eat the ice cream. Only if the respondent answers correctly (i.e., “no”) are they asked item 6 (“Why not?”) to which they should respond that it is because the ice cream contains peanut oil. If item 5 is answered incorrectly, item 6 is also marked as incorrect. The total score is the number of items correct (0–6). This value can then be converted into one of three scoring categories: high likelihood (50% or more) of limited literacy (0–1), possibility of limited literacy (2–3), and almost always indicates adequate literacy (4–6).

The NVS was developed simultaneously in English and Spanish and validated with 250 English- and 250 Spanish-speaking respondents in Tucson, Arizona [1]. Cronbach’s alpha was reported as .76 for English and .69 for Spanish. To assess criterion-related validity, NVS scores were compared to scores on the TOFHLA. Weiss et al. [1] reported significant Pearson *r* correlations (.59 English; .49, Spanish) and area under the ROC curve as .88 (95% CI, .84 – .93; *p* < .001) for English and .72 (95% CI, .66 – .79; *p* < .001) for Spanish.

1.3. Brief Health Literacy Screen (BHLS)

The BHLS⁹ is a subjective measure of health literacy consisting of three self-report items about the degree to which respondents: 1) are confident filling out medical forms independently, 2) have problems learning about a medical condition because of difficulty with written information, and 3) need help interpreting health-related materials [11]. Exact item wording and response options has varied slightly between studies. For instance, some studies specify the language of the medical forms (item 1) [12] and others do not. Typically, there are five response options that indicate frequency or intensity (e.g., ‘None of the time’ to ‘All of the time’; ‘Never’ to ‘Always’). BHLS items were originally proposed as single-item screeners because “combining questions resulted in no meaningful improvement in their screening performance of detecting inadequate or marginal health literacy” when benchmarked

against the S-TOFHLA (*N* = 332) ([2] p. 591). A second, larger validation study by Chew et al. [2] (*N* = 1796) reported that Confident with Forms had the largest AUROC curve of the three items for detecting inadequate health literacy on the S-TOFHLA (.74, 95% CI, .69 – .79) and on the REALM (.84, 95% CI, .79 – .89). Stagliano & Wallace [3] found that Confident with Forms (English) was a useful predictor of NVS scores and reported AUROC curves ranging from .79 (95% CI, .73 – .85) to .83 (95% CI, .76 – .89).

Despite Chew et al.’s (2004) recommendation to use just one item, some studies have used all three items and treated them as a scale (e.g., [13]). When all three items are used, Confident with Forms must be reverse coded in relation to the other two items.

1.4. Research context

This study of health literacy measures is drawn from a larger study, Washington Heights/Inwood Informatics Infrastructure for Community-Centered Comparative Effectiveness Research (WICER), the aim of which was to characterize the health of these two northern Manhattan neighborhoods. The large majority of neighborhood residents have roots in the Dominican Republic. In addition to measures of health literacy, the WICER survey included varied measures such as overall health, blood pressure, height, weight, waist circumference, fruit and vegetable consumption, physical activity levels, anxiety, depression, and chronic stress [14,15]. More than 6700 respondents participated in the baseline WICER survey (2010–2013) across three types of data collection sites (Columbia Community Partnership for Health, community settings including participant households, ambulatory care clinics); of those, nearly half completed a follow-up survey approximately one year later. Data in this study were drawn from the follow-up surveys and were collected in community members’ households and other community settings and ambulatory care clinics.

2. Methods

2.1. Setting and sample

This study was approved by the Columbia University Irving Medical Center Institutional Review Board. The follow-up WICER survey was administered verbally by bilingual research staff in community settings or in private rooms in two ambulatory care clinic sites within the WICER catchment area (clinical sample). Eligibility for the WICER survey was based on age (18 or over), home zip code in Washington Heights or Inwood, and ability to speak English or Spanish. For this analysis, only respondents self-reporting Hispanic ethnicity were included. Additional criteria for the analysis were complete data on the 1-year Follow-Up survey for the health literacy measures, defined as responses to all three BHLS items and at least five of the NVS items.¹⁰

2.2. Instruments

Both the NVS and BHLS were collected. The NVS was scored as described in Section 1.2. Response options from ‘Never’ to ‘Always’ were used for all three BHLS items. BHLS items were coded such that total scores ranged from 3 to 15 with high scores indicating high subjective health literacy. We also collected demographic characteristics (age, education, health insurance, place of birth, marital status, overall self-reported health, and for those born outside the 50 US states, age at arrival in the US and years living in the US).

⁸ Terms commonly used to describe this population are Hispanic (ancestry from Spanish-speaking countries, including Spain) and Latino/Latina/Latinx (ancestry from Latin American countries). For this paper, our focus is on people who have a predominantly Spanish-speaking heritage (e.g., not Portuguese), so we will use the term Hispanic throughout.

⁹ Other names for these items include the Short Literacy Survey (SLS) [7], Single Item Literacy Screeners (SILS) [8], Set of Brief Screening Questions (SBSQ) [9] and Health Literacy Screening Question Methodologies (HLSQMS) [10].

¹⁰ Cases with only four responses were also defined as complete only if one of the missing items was item 6 AND item 5 was scored as incorrect (because then item 6 would not have been offered).

Table 1
Participant Characteristics by Language of Survey Administration.

	English (<i>n</i> = 1259)		Spanish (<i>n</i> = 1729)		Total (<i>N</i> = 2988)		<i>p</i> -value
	<i>n</i>	(%)	<i>n</i>	(%)	<i>n</i>	(%)	
Recruitment Setting							< 0.001
Household	1178	(93.6)	1386	(80.2)	2564	(86)	
Ambulatory Care Clinic	81	(6.4)	343	(19.8)	424	(14)	
Gender							0.55
Female	929	(74)	1290	(75)	2219	(74)	
Male	328	(26)	436	(25)	764	(26)	
Transgender (male to female)	1	(< 0.1)	0	(0)	1	(< 0.1)	0.42 ^a
Not reported	1	(< 0.1)	3	(0.2)	4	(< 0.1)	
Education							
Never went to school	19	(2)	35	(2)	54	(2)	0.19 ^b
Eighth grade or less	430	(34)	590	(34)	1020	(34)	
Some high school	177	(14)	305	(18)	482	(16)	
High school diploma	315	(25)	398	(23)	713	(24)	0.02 ^c
Some college or technical/vocational	149	(12)	198	(12)	347	(12)	
Associate's degree	60	(5)	72	(4)	132	(4)	
Bachelor's degree	83	(7)	103	(6)	186	(6)	0.21
Master's degree	13	(1)	11	(0.6)	24	(1)	
Doctoral degree	1	(0.1)	2	(0.1)	3	(0.1)	
Not reported	12	(1)	15	(1)	27	(1)	0.82 ^d
Health Insurance							
Medicare & Medicaid	256	(20)	303	(18)	559	(19)	
Medicaid	671	(53)	926	(54)	1597	(53)	0.21
Medicare	53	(4)	82	(5)	135	(5)	
Private	72	(6)	73	(4)	145	(5)	
Other	38	(3)	58	(3)	96	(3)	0.21
None	89	(7)	120	(7)	209	(7)	
Not reported	80	(6)	167	(10)	247	(8)	
Place of Birth							0.02 ^c
Dominican Republic	966	(77)	1377	(80)	2343	(78)	
United States	194	(15)	206	(12)	400	(13)	
Argentina	1	(0.1)	1	(0.1)	2	(< 0.1)	0.21
Colombia	12	(1)	10	(0.6)	22	(0.7)	
Cuba	12	(1)	14	(0.8)	26	(0.9)	
Ecuador	13	(1)	29	(2)	42	(1)	0.21
El Salvador	3	(0.2)	9	(0.5)	12	(0.4)	
Guatemala	1	(0.1)	3	(0.2)	4	(0.1)	
Haiti	0	(0)	1	(0.01)	1	(< 0.1)	0.21
Honduras	2	(0.2)	2	(0.1)	4	(0.1)	
Mexico	10	(0.8)	25	(1)	35	(1)	
Nicaragua	3	(0.2)	4	(0.2)	7	(0.2)	0.21
Peru	2	(0.2)	2	(0.1)	4	(0.1)	
Puerto Rico	21	(2)	34	(2)	55	(2)	
Venezuela	2	(0.2)	2	(0.1)	4	(0.1)	0.21
Not reported	17	(1)	10	(0.6)	27	(1)	
Marital Status							0.21
Single, never married	373	(30)	555	(32)	928	(31)	
Married	391	(31)	515	(30)	906	(30)	
Divorced	297	(24)	351	(20)	648	(22)	0.82 ^d
Widowed	74	(6)	113	(7)	187	(6)	
Living with partner	47	(4)	60	(4)	107	(4)	
Not reported	77	(6)	135	(8)	212	(7)	0.82 ^d
Would you say that in general your health is...?							

Table 1 (continued)

	English (<i>n</i> = 1259)		Spanish (<i>n</i> = 1729)		Total (<i>N</i> = 2988)		<i>p</i> -value
	<i>n</i>	(%)	<i>n</i>	(%)	<i>n</i>	(%)	
Poor	31	(3)	26	(2)	57	(2)	< 0.001
Fair	267	(21)	361	(21)	628	(21)	
Good	291	(23)	379	(22)	670	(22)	
Very Good	250	(20)	410	(24)	660	(22)	0.55
Excellent	343	(27)	382	(22)	725	(24)	
Not reported	77	(6)	171	(10)	248	(8)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	0.42 ^a
Mean age	53	16	53	16	53	16	
Age at arrival in the United States	29 ^c	13 ^c	29 ^f	14 ^f	29 ^g	14 ^g	
Years living in the United States	28 ^e	14 ^e	27 ^f	14 ^f	28 ^g	14 ^g	0.02 ^c

Note.

^a 8th grade or less, some high school or diploma, more than high school;^b Medicaid, Medicare, dually eligible, all others; ^c Dominican Republic, United States, all others;^d Poor or Fair vs. Good or Very Good or Excellent; ^e *n* = 960; ^f *n* = 1410; ^g *n* = 2370

2.3. Procedures

Informed consent was obtained from all respondents. Respondents were recruited via probability and snowball sampling for the community sample and convenience sampling for the clinical sample. Survey questions were administered, interview style, by bilingual study staff in either English or Spanish, per respondent preference. Administration of the complete WICER survey lasted about 45–75 min. Respondents received \$25 in grocery coupons, transit tickets, or movie vouchers for their time.

2.4. Analysis

Data were cleaned through automated and manual measures using REDCap [16] and SAS 9.4. We used SPSS Statistics 27 for analysis. We calculated descriptive statistics. Cronbach's alpha was calculated to evaluate internal consistency of NVS and BHLS scores. Mean, standard deviation, corrected-item total correlation, and alpha if item deleted were calculated for all NVS and BHLS items. We examined bivariate relationships with Spearman correlations. Statistical significance was set at *p* < .001. Demographic variables that are known correlates of health literacy [17] and had bivariate correlations > |.2| were entered into multiple linear regression models to explore predictors of NVS and BHLS scores. We conducted an area under receiver operating characteristic (AUROC) curve analysis to examine the ability of the BHLS to discriminate between NVS scoring categories.

3. Results

3.1. Participants

Participant characteristics are summarized in Table 1. Of the total respondent pool (*N* = 2988), 58% responded in Spanish (*n* = 1729) and 42% in English (*n* = 1259). The mean age was 53 years (*SD* = 16). Thirty-four percent described their educational attainment as 8th grade or less, 16% had some high school, 24% had a high school diploma, and 12% had some college or technical/vocational training. For health insurance, 53% had Medicaid, 5% had Medicare, and 19% had both. Most participants were born in the Dominican Republic (78%) and 13% were born in the United States. For all those born outside the US, mean age of arrival was 29 (*SD* = 14) and mean number of years living in the US was 28 (*SD* = 14).

Table 2
Scale and Item Statistics for Newest Vital Sign and Brief Health Literacy Screener by Language.

Item	English					Spanish				
	α	<i>M</i>	<i>SD</i>	C-ITC	α if item deleted	α	<i>M</i>	<i>SD</i>	C-ITC	α if item deleted
NVS	.843 ^a					.846 ^b				
1		0.24	0.43	.523	.836		0.18	0.38	.536	.837
2		0.30	0.46	.691	.803		0.29	0.45	.741	.797
3		0.22	0.41	.585	.825		0.13	0.34	.493	.844
4		0.24	0.43	.631	.816		0.17	0.38	.554	.834
5		0.27	0.44	.658	.810		0.28	0.45	.735	.798
6		0.21	0.41	.647	.813		0.23	0.42	.705	.805
Scale		1.48	1.93				1.28	1.83		
BHLS	.797 ^c					.846 ^d				
1		3.52	1.12	.548	.814		3.38	1.13	.655	.843
2		3.34	1.38	.669	.697		3.09	1.34	.710	.793
3		3.71	1.23	.723	.634		3.37	1.31	.792	.706
Scale		10.57	3.16				9.84	3.31		

Note. C-ITC = corrected item-total correlation. Cases with missing data (81 for NVS English; 112 for NVS Spanish) were excluded from analysis.

^a $n = 1178$.

^b $n = 1617$. ^c $n = 1259$. ^d $n = 1729$.

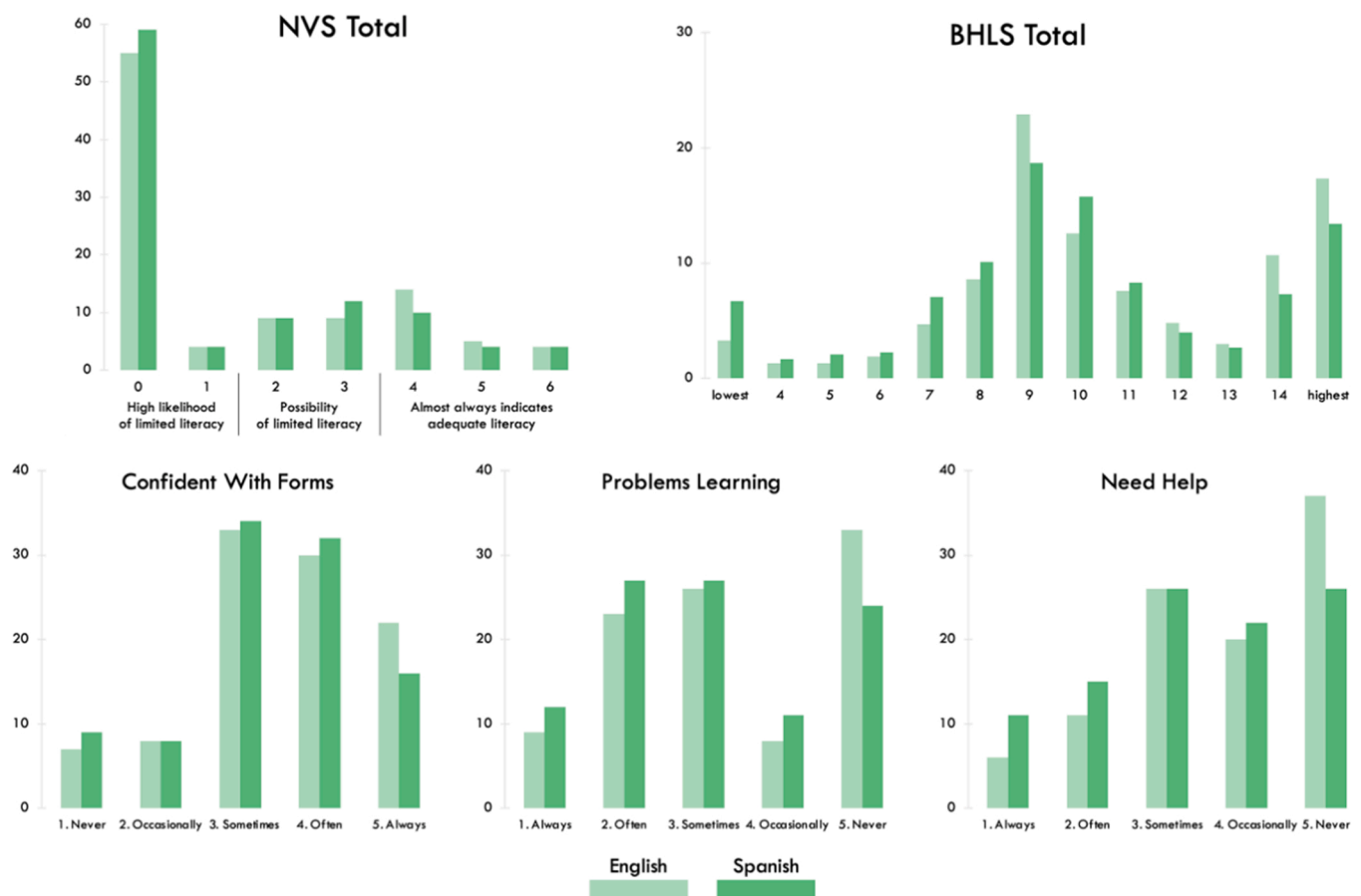


Fig. 1. Score Distributions for Total NVS Scores, Total BHLS scores, and Individual BHLS Items.

Note. The weak correlation between NVS total scores (top left) and BHLS total scores (top right) is explained by the dissimilarity in their distributions. With few exceptions, scores for all measures were fairly consistent across languages. Results shown include NVS scores with up to one item missing ($n = 81$, English; $n = 112$, Spanish). Of those with a total NVS score of zero, one item was missing for $n = 11$ for English and $n = 34$ for Spanish. There were no missing data for BHLS scores.

3.2. NVS scale and item analyses

Scale and item statistics for the NVS are shown in Table 2. For English, Cronbach's alpha was .843, scale mean was 1.48 ($SD = 1.93$) and item means ranged from 0.21 to 0.30. For Spanish, Cronbach's alpha was .846, scale mean was 1.28 ($SD = 1.83$) and item means ranged from 0.13 to 0.29. All items had strong corrected item-total correlations (.523

– .691, English; .493 – .741, Spanish).

3.3. BHLS scale and item analyses

Scale and item statistics for the BHLS are shown in Table 2. For English, Cronbach's alpha was .797, scale mean was 10.57 ($SD = 3.16$) and item means ranged from 3.34 (item 2) to 3.71 (item 3). For Spanish,

Table 3
Score distributions for NVS and BHLS.

	English (n = 1259)	Spanish (n = 1729)	Total (N = 2988)
	n (%)	n (%)	n (%)
NVS category and total score			
High likelihood of limited literacy	741 (59)	1079 (62)	1820 (61)
0	692 (55)	1015 (59)	1707 (57)
1	49 (4)	64 (4)	113 (4)
Possibility of limited literacy	230 (18)	353 (20)	583 (20)
2	116 (9)	155 (9)	271 (9)
3	114 (9)	198 (12)	312 (10)
Almost always indicates adequate literacy	288 (23)	297 (17)	585 (20)
4	177 (14)	171 (10)	348 (12)
5	58 (5)	66 (4)	124 (4)
6	53 (4)	60 (4)	113 (4)
1. How confident are you filling out medical forms by yourself / ¿Qué tan seguro(a) se siente al llenar formas usted solo(a)?			
(1) Never / Nada	85 (7)	158 (9)	243 (8)
(2) Occasionally / Un poco	104 (8)	142 (8)	246 (8)
(3) Sometimes / Algo	418 (33)	594 (34)	1012 (34)
(4) Often / Bastante	375 (30)	559 (32)	934 (31)
(5) Always / Extremadamente	277 (22)	276 (16)	553 (19)
2. How often do you have problems learning about your medical condition because of difficulty understanding written information? / ¿Qué tan seguido tiene problemas aprendiendo sobre su condición médica porque es difícil de entender información escrita?			
(1) Always / Siempre	113 (9)	204 (12)	317 (11)
(2) Often / A menudo	295 (23)	459 (27)	754 (25)
(3) Sometimes / Algunas veces	326 (26)	467 (27)	793 (27)
(4) Occasionally / Rara vez	105 (8)	184 (11)	289 (10)
(5) Never / Nunca	420 (33)	415 (24)	835 (28)
3. How often do you need to have someone help you when you read instructions, pamphlets, or other written material from your doctor or pharmacy? / ¿Qué tan seguido necesita usted tener a alguien que le ayude cuando usted lee instrucciones, folletos u otros materiales escritos por su doctor o su farmacia?			
(1) Always / Siempre	76 (6)	189 (11)	265 (9)
(2) Often / A menudo	134 (11)	257 (15)	391 (13)
(3) Sometimes / Algunas veces	334 (26)	456 (26)	790 (26)
(4) Occasionally / Rara vez	251 (20)	375 (22)	626 (21)
(5) Never / Nunca	464 (37)	452 (26)	916 (31)
BHLS Total Score			
3 (lowest literacy)	42 (3)	115 (7)	157 (5)
4	16 (1)	29 (2)	45 (2)
5	16 (1)	37 (2)	53 (2)
6	24 (2)	40 (2)	64 (2)
7	59 (5)	122 (7)	181 (6)
8	108 (9)	175 (10)	283 (9)
9	288 (23)	323 (19)	611 (20)

Table 3 (continued)

	English (n = 1259)	Spanish (n = 1729)	Total (N = 2988)
	n (%)	n (%)	n (%)
10	159 (13)	273 (16)	432 (14)
11	96 (8)	143 (8)	239 (8)
12	60 (5)	69 (4)	129 (4)
13	38 (3)	46 (3)	84 (3)
14	135 (11)	126 (7)	261 (9)
15 (highest literacy)	218 (17)	231 (13)	449 (15)

Note. Percentages may not sum to 100 because of rounding.

Cronbach's alpha was .846, scale mean was 9.84 ($SD = 3.31$) and item means ranged from 3.09 (item 2) to 3.38 (item 1). All items had strong corrected item-total correlations (.548 – .723, English; .655 – .791, Spanish). In Spanish, no item would have yielded a higher alpha if deleted. However, in English, Cronbach's alpha would increase to .814 if item 1 were deleted.

3.4. Score distributions

Score distributions for the NVS and BHLS are shown in Fig. 1 and Table 3. A bimodal distribution is visible for the NVS with 61% of respondents scoring 0 and the rest distributed normally from 1 to 6. Total BHLS scores are distributed normally from scores of 4–13 but peak again at the highest score level.

3.5. Correlation between NVS and BHLS

As shown in Table 3, the Spearman correlations between total BHLS scores and total NVS scores were statistically significant but weak for both languages ($r_s = .21, p < .001$, English; $r_s = .19, p < .001$, Spanish). The cross-tabulation of the BHLS items with NVS categories in Fig. 2 demonstrates that although the majority of respondents were in the lowest NVS category (i.e., “high likelihood of limited literacy”) their ratings on BHLS items spanned all five levels of subjective literacy.

3.6. Relationships with demographic variables

We examined age, age at arrival in the US, years since arrival in the US, education, and self-reported health for bivariate correlations $> |.2|$ with total NVS and BHLS scores. Based on these, we selected age and education for multiple regression models, shown in Table 4. Together, age and education accounted for 13% of the variance in total NVS scores in English but only 6% in Spanish ($R^2 = .13, p < .001$; $R^2 = .06, p < .001$). Age and education accounted for 14% (English) and 15% (Spanish) of the variance in total BHLS scores ($R^2 = .14, p < .001$; $R^2 = .15, p < .001$). In all four models, controlling for education decreased the impact of age.

3.7. Area under the receiver operating characteristic curve (AUROC)

For BHLS total scores, the AUROC curves were .606 (English) and .605 (Spanish) for discriminating the lowest NVS lowest scoring category, as shown in Fig. 3 and Table 5.

4. Discussion and conclusion

4.1. Discussion

In this study, scores for both the NVS and BHLS demonstrated good internal consistency but were weakly correlated with each other.

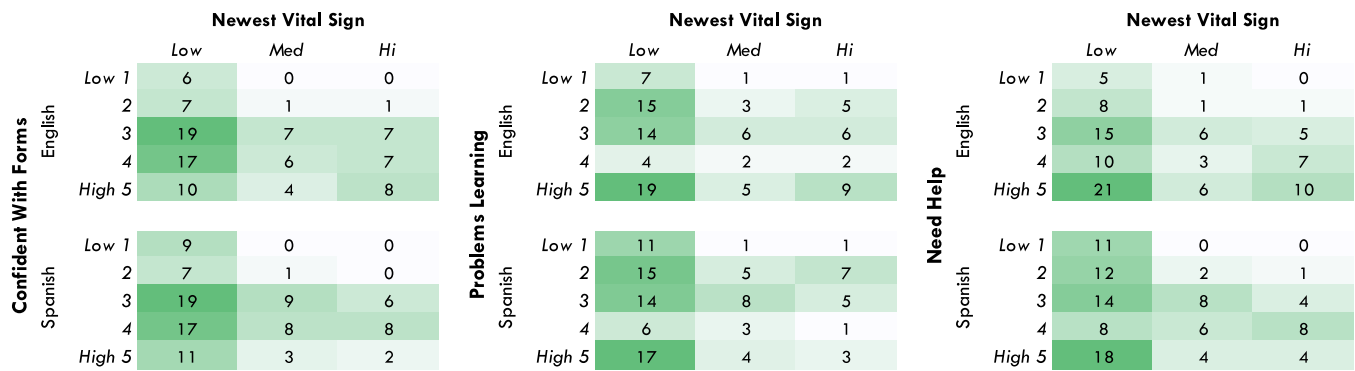


Fig. 2. Cross-Tabulation Heat Map of NVS Categories and BHLS Item Score Levels by Language of Administration.

Note. English $n = 1259$ (top row); Spanish $n = 1729$ (bottom row). Values for each 3×5 matrix sum to 100 and represent the proportion of respondents who fell within that cell. Darker shades represent a larger proportion of participants. This heat map demonstrates that although the majority of respondents were in the lowest NVS category (i.e., lowest level of objective health literacy) their ratings on BHLS items nevertheless spanned all five levels of subjective literacy.

Table 4

Multiple Regressions of Associations Between Demographic Variables and Health Literacy Scores.

Variable	Total NVS Score, English			Total NVS Score, Spanish			Total BHLS Score, English			Total BHLS Score, Spanish		
	B	SE	β	B	SE	β	B	SE	β	B	SE	β
Constant	1.19	.28		1.99	.24		9.75	.45		8.91	.40	
Age	-.02	.00	-0.13	-.02	.00	-0.18	-.02	.00	-0.12	-.03	.01	-0.13
Education	0.33	.04	0.28	0.12	.03	0.10	0.58	.06	0.30	0.66	.05	0.31
R	0.36			0.24			0.37			0.39		
R ²	.13			0.06			0.14			0.15		

* $p < .001$

Reference: NVS Scores 0 - 1 (high likelihood of limited health literacy)

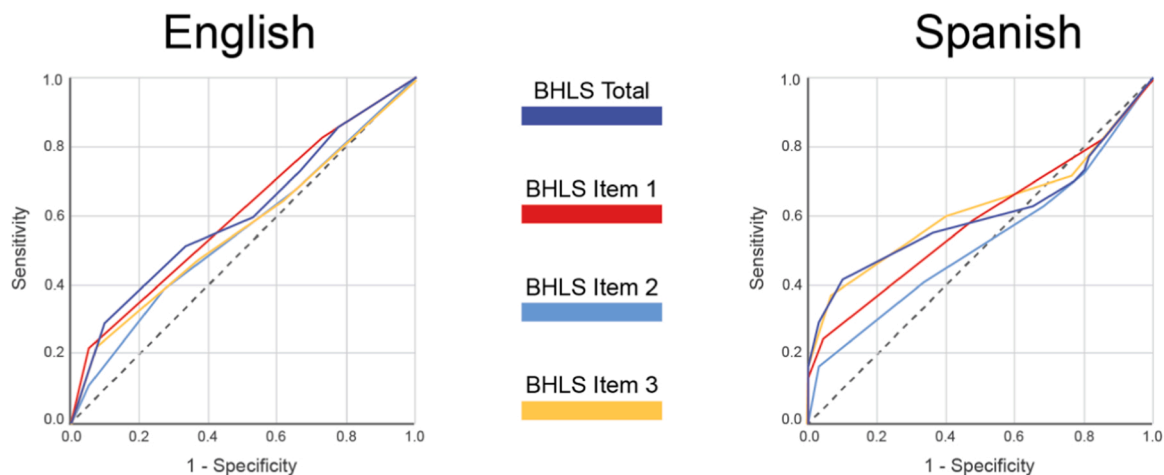


Fig. 3. Area Under the Receiver Operating Characteristic (AUROC) Curves.

Although the correlations were statistically significant, they were likely driven by the sample size ($N = 2988$). These results are in contrast to Stagliano and Wallace [3] who found that BHLS scores were good predictors of NVS scores in English among non-Hispanic participants with higher educational attainment than our sample had. Of note in our sample, although the majority of respondents were in the lowest NVS

category, their ratings on BHLS items spanned all five levels of subjective literacy. This suggests that individuals may overestimate their literacy or perhaps that the numeracy requirements to score high on the NVS are different than what is needed to complete typical tasks related to health care.

Our results are consistent with Waters et al. [18], who examined the

Table 5

Area Under the Receiver Operating Characteristic (AUROC) Curves.

Test Result Variables	NVS High Likelihood of Limited Literacy	
	English	Spanish
BHLS Total	.606	.605
BHLS Item 1	.611	.586
BHLS Item 2	.558	.523
BHLS Item 3	.566	.617

English NVS and BHLS scores of 1005 respondents (9.7% Hispanic) as indicators of latent factors (objective and subjective literacy, respectively) and found a correlation of only .18 ($p < .001$) between the factors. They are also similar to those of Gutierrez and colleagues [19] who surveyed 265 patients in a public clinic in Dallas, TX (62% Hispanic, 54% in Spanish) and found that 53% rated Confidence with Forms highly but only 18.9% had NVS scores indicating adequate literacy.

Other studies with Spanish-speaking cohorts reported somewhat better performance by Confident with Forms against criterion measures than we found. Hadden et al. [12] surveyed 200 Spanish-speaking participants using the NVS as a benchmark and two versions of Confident with Forms that specified the language of the forms. They found that Confident with English Forms had an AUROC curve of .70 whereas it was .63 (confidence intervals were not reported) for Confident with Spanish Forms. Cordasco et al. [20] studied a geriatric, monolingual Spanish-speaking cohort ($N = 134$) and found that although Confident with Forms reasonably discriminated S-TOFHLA scores (AUROC .82, 95% CI. 75 – .89) educational attainment was slightly better (AUROC .88, 95% CI. 78 – .97); the other two BHLS items had AUROC curves less than .50.

Given the limitations of our study design, we can only speculate as to what accounts for the weak relationship between the BHLS and NVS in both English and Spanish for Hispanic participants. We propose several potential explanations. First, NVS items were of very high difficulty for this population and demonstrated floor effects, with 61% of all respondents scoring 0. Perhaps a lower-difficulty measure might have produced more score variability and potentially a stronger correlation with BHLS scores. However, the TOFHLA has lower difficulty than the NVS [21,22] and studies in the US that compared the English BHLS or its items to TOFHLA or S-TOFHLA scores have found widely varying correlations ($r_s = .21$ to $.49$) and AUROC curves (.62 – .87), so no clear pattern emerges [7,23,24].

A second potential explanation is that BHLS scores could have been inflated due to social desirability bias or overconfidence on the part of the respondents. Health care providers tend to overestimate patients' literacy [25,26], so it is reasonable to question the accuracy of individuals' assessment of their own abilities [27] and their subsequent confidence in perceived abilities. These potential pitfalls are inherent limitations of subjective measures and may not be easily remedied.

Lastly, it is possible that the problem is conceptual. The premise underlying comparison of the BHLS to the NVS is that subjective and objective health literacy (with an emphasis on numeracy) are closely related concepts. If the NVS and BHLS both have good construct validity, our study calls that premise into question and raises concerns about the utility of these concepts. In light of the many and varied definitions of health literacy and in the varying ways those definitions have been operationalized into measurement, there has been a call for an improved definition of health literacy that better acknowledges its contextual nature [28] and its importance as a social determinant of health [29]. The Department of Health and Human Services (DHHS) recently answered this call with updated definitions for *personal health literacy* and *organizational health literacy* [30]. This change warrants a re-evaluation of health literacy measurement methods to ensure they are keeping pace with evolving definitions.

4.1.1. Strengths and limitations

A notable strength of this study is the very large sample size of our cohort ($N = 2988$) and the fact that the English and Spanish groups were very similar along demographic characteristics. Limitations include that it was a convenience sample with a disproportionate number of women. Also, the wording of the BHLS items did not specify whether or not the forms, written materials, etc. were in the respondent's preferred language. Our participants were overwhelmingly from or had roots in the Dominican Republic; results may not be applicable to Hispanic people from other regions.

4.2. Conclusion

As a scale, the BHLS had good internal consistency in both English and Spanish but had poor ability to discriminate between NVS score categories in a large, urban Hispanic cohort. The NVS also functioned well in both languages and had the advantage of sidestepping the potential pitfalls associated with subjective self-report measures. However, the NVS was a very high difficulty measure in this population which may make it clinically useful for identifying people at risk of inadequate health literacy, but yields score distributions with strong floor effects that preclude discrimination at low levels of the trait continuum, and thus limiting its utility for some research studies.

4.3. Practice implications

Measures of subjective health literacy may fail to identify individuals with limited literacy and so objective measures may be preferable. Researchers desiring score variability within lower levels of health literacy may want to use objective measures that are of lower difficulty than the NVS.

CRedit authorship contribution statement

Adriana Arcia: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. **Anthony T. Pho:** Methodology, Data curation, Writing – review & editing. **Maichou Lor:** Data curation, Writing – original draft, Writing – review & editing. **Suzanne Bakken:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition.

Acknowledgments

This study was funded by the Agency for Healthcare Research and Quality [R01HS019853, R01HS022961]; the New York State Department of Economic Development NYSTAR [C090157]. Maichou Lor received postdoctoral fellowship funding from the National Institute for Nursing Research [T32NR007969]. Anthony Pho received pre-doctoral funding from the Robert Wood Johnson Foundation Future of Nursing Scholars Program. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors are grateful to Niurka Suero-Tejeda for her contributions to project management and data collection.

Conflicts of Interest

We have no known conflicts of interest to disclose.

References

- [1] Weiss BD, Mays MZ, Martz W, Castro KM, DeWalt DA, Pignone MP, et al. Quick assessment of literacy in primary care: the Newest Vital Sign. *Ann Fam Med* 2005; 3:514–22. <https://doi.org/10.1370/afm.405>.
- [2] Chew LD, Griffin JM, Partin MR, Noorbaloochi S, Grill JP, Snyder A, et al. Validation of screening questions for limited health literacy in a large VA outpatient population. *J Gen Intern Med* 2008;23:561–6. <https://doi.org/10.1007/s11606-008-0520-5>.

- [3] Stagliano V, Wallace LS. Brief health literacy screening items predict Newest Vital Sign scores. *J Am Board Fam Med* 2013;26:558–65. <https://doi.org/10.3122/jabfm.2013.05.130096>.
- [4] Mantwill S, Allam A, Camerini A-L, Schulz PJ. Validity of three brief health literacy screeners to measure functional health literacy - evidence from five different countries. *J Health Commun* 2018;23:153–61. <https://doi.org/10.1080/10810730.2017.1417515>.
- [5] Ramirez-Zohfeld V, Rademaker AW, Dolan NC, Ferreira MR, Eder MM, Liu D, et al. Comparing the performance of the S-TOFHLA and NVS among and between English and Spanish speakers. *J Health Commun* 2015;20:1458–64. <https://doi.org/10.1080/10810730.2015.1018629>.
- [6] U.S. Census Bureau. QuickFacts: United States 2019. <https://www.census.gov/quickfacts/fact/table/US/PST045219> (accessed March 15, 2021).
- [7] Dageforde LA, Cavanaugh KL, Moore DE, Harms K, Wright A, Pinson CW, et al. Validation of the written administration of the short literacy survey. *J Health Commun* 2015;20:835–42. <https://doi.org/10.1080/10810730.2015.1018572>.
- [8] Quinzanos I, Hirsh JM, Bright C, Caplan L. Cross-sectional correlation of single-item health literacy screening questions with established measures of health literacy in patients with rheumatoid arthritis. *Rheuma Int* 2015;35:1497–502. <https://doi.org/10.1007/s00296-015-3238-9>.
- [9] van Schaik TM, Jørstad HT, Twickler TB, Peters RJG, Tijssen JPG, Essink-Bot ML, et al. Cardiovascular disease risk and secondary prevention of cardiovascular disease among patients with low health literacy. *Neth Heart J* 2017;25:446–54. <https://doi.org/10.1007/s12471-017-0963-6>.
- [10] Collins SA, Currie LM, Bakken S, Vawdrey DK, Stone PW. Health literacy screening instruments for eHealth applications: a systematic review. *J Biomed Inf* 2012;45: 598–607. <https://doi.org/10.1016/j.jbi.2012.04.001>.
- [11] Chew LD, Bradley KA, Boyko EJ. Brief questions to identify patients with inadequate health literacy. *Fam Med* 2004;36:588–94.
- [12] Hadden KB, Prince LY, Rojo MO, Selig JP, McElfish PA. Screening patients who speak spanish for low health literacy. *Health Lit Res Pr* 2019;3:e110–6. <https://doi.org/10.3928/24748307-20190408-03>.
- [13] McNaughton C, Wallston KA, Rothman RL, Marcovitz DE, Storrow ABShort. Subjective measures of numeracy and general health literacy in an adult emergency department. *Acad Emerg Med* 2011;18:1148–55. <https://doi.org/10.1111/j.1553-2712.2011.01210.x>.
- [14] Arcia A, Bales ME, Brown W, Co MC, Gilmore M, Lee YJ, et al. Method for the development of data visualizations for community members with varying levels of health literacy. *AMIA Annu Symp Proc* 2013;2013:51–60.
- [15] Arcia A, Suero-Tejeda N, Bales ME, Merrill JA, Yoon S, Woolen J, et al. Sometimes more is more: iterative participatory design of infographics for engagement of community members with varying levels of health literacy. *J Am Med Inf Assoc* 2016;23:174–83. <https://doi.org/10.1093/jamia/ocv079>.
- [16] Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inf* 2009;42: 377–81. <https://doi.org/10.1016/j.jbi.2008.08.010>.
- [17] Shah LC, West P, Bremmeyr K, Savoy-Moore RT. Health literacy instrument in family medicine: the “Newest Vital Sign” ease of use and correlates. *J Am Board Fam Med* 2010;23:195–203. <https://doi.org/10.3122/jabfm.2010.02.070278>.
- [18] Waters EA, Biddle C, Kaphingst KA, Schofield E, Kiviniemi MT, Orom H, et al. Examining the interrelations among objective and subjective health literacy and numeracy and their associations with health knowledge. *J Gen Intern Med* 2018; 33:1945–53. <https://doi.org/10.1007/s11606-018-4624-2>.
- [19] Gutierrez N, Kindratt TB, Pagels P, Foster B, Gimpel NE. Health literacy, health information seeking behaviors and internet use among patients attending a private and public clinic in the same geographic area. *J Community Health* 2014;39:83–9. <https://doi.org/10.1007/s10900-013-9742-5>.
- [20] Cordasco KM, Homeier DC, Franco I, Wang P-C, Sarkisian CA. Health literacy screening of geriatric monolingual Spanish-speaking patients using single-item literacy screening questions and education. *Health Educ J* 2012;71:597–605. <https://doi.org/10.1177/0017896911411764>.
- [21] Morrison AK, Schapira MM, Hoffmann RG, Brousseau DC. Measuring health literacy in caregivers of children: a comparison of the Newest Vital Sign and S-TOFHLA. *Clin Pediatr (Philo)* 2014;53:1264–70. <https://doi.org/10.1177/0009922814541674>.
- [22] Barber MN, Staples M, Osborne RH, Clerehan R, Elder C, Buchbinder R. Up to a quarter of the Australian population may have suboptimal health literacy depending upon the measurement tool: results from a population-based survey. *Health Promot Int* 2009;24:252–61. <https://doi.org/10.1093/heapro/dap022>.
- [23] Kaphingst KA, Goodman MS, MacMillan WD, Carpenter CR, Griffey RT. Effect of cognitive dysfunction on the relationship between age and health literacy. *Patient Educ Couns* 2014;95:218–25. <https://doi.org/10.1016/j.pec.2014.02.005>.
- [24] Cavanaugh KL, Osborn CY, Tentori F, Rothman RL, Ikizler TA, Wallston KA. Performance of a brief survey to assess health literacy in patients receiving hemodialysis. *Clin Kidney J* 2015;8:462–8. <https://doi.org/10.1093/ckj/sfv037>.
- [25] Goggins K, Wallston KA, Mion L, Cawthon C, Kripalani S. What patient characteristics influence nurses’ assessment of health literacy? *J Health Commun* 2016;21:105–8. <https://doi.org/10.1080/10810730.2016.1193919>.
- [26] Dickens C, Lambert BL, Cromwell T, Piano MR. Nurse overestimation of patients’ health literacy. *J Health Commun* 2013;18(Suppl 1):62–9. <https://doi.org/10.1080/10810730.2013.825670>.
- [27] Kruger J, Dunning D. Unskilled and unaware of it: how difficulties in recognizing one’s own incompetence lead to inflated self-assessments. *J Pers Soc Psychol* 1999; 77:1121–34. <https://doi.org/10.1037/0022-3514.77.6.1121>.
- [28] Pleasant A, Rudd RR, O’Leary C, Paasche-Orlow MK, Allen MP, Alvarado-Little W, et al. Considerations for a New Definition of Health Literacy. Washington, DC: Natl Acad Med 2016.
- [29] Nutbeam D, Lloyd JE. Understanding and responding to health literacy as a social determinant of health. *Annu Rev Public Health* 2021;42:159–73. <https://doi.org/10.1146/annurev-publhealth-090419-102529>.
- [30] Santana S, Brach C, Harris L, Ochiai E, Blakey C, Bevington F, et al. Updating health literacy for healthy people 2030: defining its importance for a new decade in public health. *J Public Health Manag Pr* 2021;27:S258–64. <https://doi.org/10.1097/PHH.0000000000001324>.