Original Research

Health Literacy, Education, and Internal Consistency of Psychological Scales

Xuewei Chen, PhD; Elizabeth Schofield, MPH; Heather Orom, PhD; Jennifer L. Hay, PhD; Marc T. Kiviniemi, PhD, CPH; and Erika A. Waters, PhD

ABSTRACT

Background: Measurement error might lead to biased estimates, causing ineffective interventions and service delivery. Identifying measurement error of health-related instruments helps develop accurate assessment of health-related constructs. **Objective:** We compared the internal consistency of eight psychological scales used in health research in groups with adequate versus limited health literacy and in groups with higher versus lower education. **Methods:** Participants (N = 1,005) from a nationally representative internet panel completed eight self-report scales: (1) information avoidance, (2) cognitive causation, (3) unpredictability, (4) perceived severity, (5) time orientation, (6) internal health locus of control, (7) need for cognition, and (8) social desirability. The first four assess beliefs about diabetes and colon cancer. We used the Newest Vital Sign to categorize participants' health literacy (limited vs. adequate). We also categorized participants' education (high school or less vs. more than high school). We compared the Cronbach's alpha for each psychological scale between groups with different health literacy and education levels using the Feldt test. **Key Results:** Among all the 13 subscales, scale internal consistency was significantly lower among people with limited health literacy than those with adequate health literacy for five subscales: information avoidance for colon cancer (0.80 vs. 0.88), unpredictability of diabetes (0.84 vs. 0.88), perceived severity for diabetes (0.66 vs. 0.75), need for cognition (0.63 vs. 0.82), and social desirability (0.52 vs. 0.68). Internal consistency was significantly lower among people who had a high school education or less than among those with more than a high school education for four scales: perceived severity of diabetes (0.70 vs. 0.75), present orientation (0.60 vs. 0.66), need for cognition (0.73 vs. 0.80), and social desirability (0.61 vs. 0.70). **Conclusions:** Several psychological instruments demonstrated significantly lower internal consistency when used in a sample with limited health literacy or education. To advance health disparities research, we need to develop new scales with alternative conceptualizations of the constructs to produce a measure that is reliable among multiple populations. **[*HLRP: Health Literacy Research and Practice*. 2021;5(3):e244-e255.]**

Plain Language Summary: We compared the internal consistency of several psychological scales in groups with adequate versus limited health literacy and higher versus lower education. For several scales, internal consistency was significantly lower among (1) people with limited health literacy compared those who have adequate health literacy and/or (2) people who had a high school education or less compared to those with more than a high school education.

Psychological measurement instruments can be used to better understand psychological phenomena, explicate the psychological processes that make an intervention successful (or not), help researchers test hypotheses, identify target populations through conducting needs assessments, and screen people for clinical services (Sturm & Ash, 2005). However, to maximize their utility for advancing theory and improving interventions, psychological measurement instruments must demonstrate that they produce scores that are valid and reliable across samples. Importantly, reliability is a necessary (although not itself sufficient) criterion for validity (Crocker & Algina, 1986; Onwuegbuzie & Daniel, 2002).

Internal consistency is one way of assessing reliability (Henson & Thompson, 2002). A measure yielding scores with low internal consistency will add error variance to a statistical model and, therefore, could lead to incorrect conclusions that a construct and outcome are more weakly related to each other than is actually the case (Reinhardt, 1991). This may

also lead to the erroneous inclusion or exclusion of constructs in intervention models or of intervention components that target the latent construct, thereby inhibiting intervention effects in behavioral intervention trials or leading to inaccurate behavioral theories.

INDIVIDUAL DIFFERENCES IN INTERNAL CONSISTENCY IN PSYCHOLOGICAL MEASUREMENT INSTRUMENTS

Scale internal consistency is not stable across populations with varying characteristics (Henson et al., 2001). However, many psychological measures are still used with participants who are dramatically different (in terms of demographics, health literacy, and other critical factors) from the sample within which the scale was originally developed. These differences in the characteristics of subsequent populations may affect the internal consistency of test scores across groups.

Many instruments have been developed with samples of college students. Traditionally, these students have had higher socioeconomic status, literacy skills, and have been less demographically diverse than the general population (Hanel & Vione, 2016). Instruments developed in samples of college students may be interpreted differently or have less personal relevance when used in samples of non-college students, which might reduce the scale's internal consistency (Shepperd et al., 2016).

There is evidence that internal consistencies for psychological instruments differ across diverse groups of study participants who vary according to language, cultural background, education, and reading skills (Gjersing et al., 2010; Shepperd et al., 2016; Taras et al., 2009). For example, the internal consistency for three key psychological measures—the behavioral inhibition scale/behavioral activation scale (BIS/BAS), the regulatory focus questionnaire (RFQ), and the need for cognition scale (NCS)—differed between groups with high and low education (Shepperd et al., 2016). The BIS/BAS, RFQ, and NCS were all originally developed using college students (Carver & White, 1994; Cohen, 1957; Higgins et al., 2001). Shepperd et al. (2016) found that these three measures all have lower internal consistency for people with a high school education or less, and higher internal consistency for people with more than a high school education. They attribute lower internal consistency among people with lower education to their having lower literacy skills and being infrequent readers.

Limited internal consistency in the BIS/BAS, RFQ, and NCS suggests that other scales used in health behavior research may have similar limitations. For example, avoidance of health information (hereafter "information avoidance"), cognitive causation (i.e., the belief that thinking about a health problem will cause it to occur), beliefs that disease onset is unpredictable, perceived severity of disease, time orientation (i.e., the extent to which people think about the future or present), internal health locus of control (i.e., the extent to which people believe that their health is in their control), and social desirability (i.e., the tendency to answer questions in a way one thinks others would approve of) have all been used in health research. However, to our knowledge, with the exception of NCS, the potential for differential reliability of these scales as a function of health literacy or education has not been previously examined.

HEALTH LITERACY

Health literacy could also affect internal consistency. Health literacy represents the ability to "obtain, process, and

Xuewei Chen, PhD, is an Assistant Professor, School of Community Health Sciences, Counseling and Counseling Psychology, Oklahoma State University. Elizabeth Schofield, MPH, is a Research Biostatistician, Department of Psychiatry & Behavioral Sciences, Memorial Sloan-Kettering Cancer Center. Heather Orom, PhD, is an Associate Professor, Department of Community Health and Health Behavior, University at Buffalo. Jennifer L. Hay, PhD, is an Attending Psychologist, Department of Psychiatry & Behavioral Sciences, Memorial Sloan-Kettering Cancer Center. Marc T. Kiviniemi, PhD, CPH, is a Professor and the Chair, Department of Health, Behavior, and Society, University of Kentucky. Erika A. Waters, PhD, is a Professor, Department of Surgery (Public Health Sciences Division), Washington University School of Medicine in St. Louis.

Address correspondence to Xuewei Chen, PhD, School of Community Health Sciences, Counseling and Counseling Psychology, Oklahoma State University, 429 Willard Hall, Stillwater, OK 74078; email: xuewei.chen@okstate.edu.

Grant: This study was supported by a grant (R01CA197351) from the National Cancer Institute.

Disclosure: The authors have no relevant financial relationships to disclose.

Received: April 28, 2020; Accepted: September 15, 2020

doi: 10.3928/24748307-20210728-01

understand basic health information and services needed to make appropriate health decisions" (Kindig et al., 2004). Reading, writing, and numeracy skills are essential components of health literacy (Parker et al., 1995). Besides education, other factors may influence people's health literacy, including living in poverty, race/ethnicity, age, and disability (Kutner et al., 2006). Thus, comparing groups with different health literacy levels contributes novel findings to the literature.

Survey items that ask about health are potentially affected by people's ability to process and understand health information. Therefore, it is likely that an instrument that has good internal consistency among people with adequate health literacy may have poor internal consistency among those with limited health literacy. Given the high levels of limited health literacy in the United States (U.S. Department of Health and Human Services, 2008), understanding these potential distinctions in the internal consistency of some psychological scales will be practically and methodologically useful for basic social and health psychology research and applied public health promotion research. For example, identifying the psychological scales that have poor internal consistency among people with limited health literacy would lead to scale revision and construct reconceptualization to reduce measurement error so that future research would generate more accurate and less biased results.

Although researchers have studied measurement invariance by age and gender of 15 psychological measures, including NCS (Hussey & Hughes, 2020), they have not examined invariance by education or health literacy. Researchers also determined that people with limited health literacy may respond differently to purportedly validated scales than those with adequate health literacy (Taple et al., 2019), but none of the eight scales identified in the previous section have been evaluated to determine whether there are differences in their internal consistency according to health literacy.

OBJECTIVE AND HYPOTHESES

In the present study, we examined whether the internal consistency of eight psychological scales used in health research differed by health literacy or education. We contribute to the literature in three key ways: (1) using a nationally representative sample (vs. a convenience sample in Shepperd et al., 2016), (2) comparing groups with different health literacy levels as well as education levels (vs. only education), and (3) examining the internal consistency of several scales used in health research that were not included in studies by Shepperd et al. (2016) or Hussey & Hughes (2020). We hypothesize that the scores of people with lower education and lower health literacy will be less internally consistent than the scores of people with higher education and higher health literacy.

METHODS

The study was approved by the University at Buffalo Institutional Review Board. This article presents a secondary analysis of data collected for a larger study designed to examine possible psychological mechanisms underlying not knowing one's risk for common diseases (Orom et al., 2018).

Sample

GfK, a market research firm with an academic research arm, conducted recruitment and data collection from May to June 2016. GfK maintained a standing representative panel (KnowledgePanel) of 55,000 people. For this study, GfK sent email invitations to 1,818 KnowledgePanel members, and 1,033 (56.8%) of them completed the survey. Responses for 26 participants were dropped because they met two or more of the following criteria for inattentive responding: (1) completed the survey in less than 5.5 minutes (one-fourth of the median time of 22 minutes), (2) marked identical responses or straight-lined for at least one-half of the eight question grids, (3) failed both of the survey validation items (asking participants to select *somewhat agree* for one item and *somewhat disagree* for the other item), and (4) gave different answers to a repeated factual question about their health insurance status. Two additional participants were excluded due to having prevalent diabetes and colorectal cancer. Thus, we included a final sample of 1,005 participants in our data analyses.

Measure

We examined the internal consistency of eight psychological instruments that were included in the original study design (Orom et al., 2018; Waters et al., 2018). Several of the scales had subscales. See **Table 1** for information about the Flesch-Kincaid Reading Level (FKRL) for each subscale, as well as the population(s) in which it was developed.

The Newest Vital Sign (NVS) is a measure of health literacy that contains six questions to test participants' understanding of information on a mock-up nutrition label for ice cream (Weiss et al., 2005). Participants received 1 point for each correct answer. A missing response was considered incorrect and received a score of 0. The NVS score was categorized as follows: high likeli-

TABLE 1

The Eight Psychological Scales Compared in this Study

Scale	Population in Which the Scale Was Developed	Flesch-Kincaid Reading Level
Information Avoidance Scale (Howell & Shepperd, 2016) *Diabetes* 1. I feel like I have enough information to know my risk of getting diabetes. 2. I would rather not know about diabetes. 3. I would prefer to avoid learning about diabetes. 4. Even if it will upset me, I want to know about diabetes. 5. I want to know about diabetes. 6. I can think of situations in which I would rather not know about diabetes. 7. It is important to know about diabetes. 8. I want to know about diabetes immediately. *Colon cancer* 1. I feel like I have enough information to know my risk of getting colon cancer. 2. I would rather not know about colon cancer. 3. I would prefer to avoid learning about colon cancer. 4. Even if it will upset me, I want to know about colon cancer. 5. I want to know about colon cancer. 6. I can think of situations in which I would rather not know about colon cancer. 7. It is important to know about colon cancer. 8. I want to know about colon cancer immediately.	Undergraduates, high school students, U.S. adults, African Americans, and adult women	7.5 7.2
Cognitive Causation Scale (Hay et al., 2014) *Diabetes* 1. If I think too hard about the possibility of getting diabetes, I could get it. 2. If I don't believe I will get diabetes, I won't. 3. Negative thoughts about getting diabetes might make me get it. 4. Considering that I could get diabetes might bring on bad luck. 5. Too much thought about diabetes risk could encourage the disease. 6. Thinking that I am likely to get diabetes may give me diabetes. 7. In general, if a person thinks about the possibility of getting diabetes, they are more likely to get it. *Colon cancer* 1. If I think too hard about the possibility of getting colon cancer, I could get it. 2. If I don't believe I will get colon cancer, I won't. 3. Negative thoughts about getting colon cancer might make me get it. 4. Considering that I could get colon cancer might bring on bad luck. 5. Too much thought about colon cancer risk could encourage the disease. 6. Thinking that I am likely to get colon cancer may give me colon cancer. 7. In general, if a person thinks about the possibility of getting colon cancer, they are more likely to get it.	Undergraduate psychology students, community men, immigrants, African Americans, and Black Caribbeans	8.1 8.0

hood of limited health literacy (0-1 correct); possibility of limited health literacy (2-3 correct); almost always adequate health literacy (4-6 correct) (Weiss et al., 2005). In our sample, only 64 (6%) participants scored in the lowest

TABLE 1 (*continued*)

The Eight Psychological Scales Compared in this Study

Scale	Population in Which the Scale Was Developed	Flesch-Kincaid Reading Level
Unpredictability of Cancer Scale (Hay et al., 2014) *Diabetes* 1. Anybody can get diabetes, no matter what they do. 2. Diabetes can strike anyone at any time. 3. You never know who is going to get diabetes. *Colon cancer* 1. Anybody can get colon cancer, no matter what they do. 2. Colon cancer can strike anyone at any time. 3. You never know who is going to get colon cancer.	Undergraduate psychology students, community men, immigrants, African Americans, and Black Caribbeans	 6.5 6.3
Perceived Severity Scale (Moss-Morris et al., 2002) *Diabetes* 1. Diabetes is a serious condition. 2. If I had diabetes, it would have major consequences on my life. 3. If I had diabetes, it would not have much effect on my life. 4. If I had diabetes, it would have serious financial consequences. 5. If I had diabetes, it would cause difficulties for those who are close to me. *Colon cancer* 1. Colon cancer is a serious condition. 2. If I had colon cancer, it would have major consequences on my life. 3. If I had colon cancer, it would not have much effect on my life. 4. If I had colon cancer, it would have serious financial consequences. 5. If I had colon cancer, it would cause difficulties for those who are close to me.	People from seven illness groups (asthma, diabetes, rheumatoid arthritis, chronic pain, acute pain, myocardial infarction, and multiple sclerosis) in New Zealand and an HIV patient group in the United Kingdom	 7.1 7.0
Time Orientation Scale (Lukwago et al., 2001) *Present orientation* 1. My day-to-day is too busy to think about the future. 2. There's no sense in thinking about the future before it gets there. 3. What happens to me in the future is out of my control. *Future orientation* 1. I have a plan for what I want to do in the next 5 years of my life. 2. The choices I have made in life clearly show that I think about the future. 3. I often think about how my actions today will affect my health when I am older.	African American women from urban housing communities with low income	 6.1 5.6
Internal Health Locus of Control Scale (Wallston et al., 1978) 1. If I get sick, it is my own behavior which determines how soon I get well again. 2. I am in control of my health. 3. When I get sick, I am to blame. 4. The main thing which affects my health is what I myself do. 5. If I take care of myself, I can avoid illness. 6. If I take the right actions, I can stay healthy.	A sample of 90% White people and 74% having at least some college education	4.1

category (0-1). Thus, we treated health literacy as a binary variable, dividing health literacy as limited (NVS 0-3) or adequate (NVS 4-6). This approach has been used in previous studies (Ghaddar et al., 2012; Griffey et al., 2014; Hudon et al., 2012; Protheroe et al., 2017). The NVS was selected because it could be used in the context of a rela-

TABLE 1 (*continued*)

The Eight Psychological Scales Compared in this Study

Scale	Population in Which the Scale Was Developed	Flesch-Kincaid Reading Level
Need for Cognition Scale (Sherrard & Czaja, 1999) 1. I like to have the responsibility of handling a situation that requires a lot of thinking. 2. Thinking is not my idea of fun. 3. I would rather do something that requires little thought than something that is sure to challenge my thinking abilities. 4. I only think as hard as I have to. 5. I really enjoy a task that involves coming up with new solutions to problems. 6. Learning new ways to think doesn't excite me very much. 7. I prefer my life to be filled with puzzles that I must solve.	Primarily undergraduate university and technical institute students	6.9
Social Desirability Scale (Paulhus, 1991) 1. There have been occasions when I have taken advantage of someone. 2. I always obey laws, even if I'm unlikely to get caught. 3. I never swear. 4. I sometimes drive faster than the speed limit. 5. I have done things that I don't tell other people about. 6. I never take things that don't belong to me. 7. I have taken sick-leave from work or school even though I wasn't really sick. 8. I have never damaged a library book or store merchandise without reporting it.	College students	6.2

tively brief, self-administered online survey and yields reliable and valid scores among various populations across different age ranges, races/ethnicities, and health conditions (Shealy & Threatt, 2016; Weiss et al., 2005).

Demographic variables included sex, age, race/ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, and race/ethnicity unknown), and education (less than high school, high school, some college, Bachelor's degree or higher).

Data Analysis

For each measure, we computed internal consistency using Cronbach's alpha overall, by health literacy group, and by education level. To compare alphas across health literacy and education levels, we use the Feldt test (Feldt, 1969). The Feldt test can be used to compare alphas across two groups, where the statistic W, is the ratio of (1-alpha2) to (1-alpha1) and follows an F distribution with degrees of freedom v1=N1–1 and v2=N2–1. The sample of 1,000 participants exceeds the minimum requirements for the Feldt test (Feldt & Kim, 2006). We collapsed the education variable into two levels (high school and below vs. above high school) because the Feldt test cannot be used to compare alphas across more than two groups. See **Table A** for alphas across the four education levels.

RESULTS

About 20% of our sample had limited health literacy ($n = 206$) and about 37% ($n = 372$) of the sample had completed high school or less. Just 14% ($n = 136$) had both limited health literacy and low formal education, and 56% ($n = 563$) had adequate health literacy and high formal education. However, among people with low education, 37% had limited health literacy; among people with high education, only 11% had limited health literacy. Our dichotomous variables of health literacy and education were correlated (phi = 0.31, $p < .001$). See **Table 2** for detailed information about sample characteristics.

Health Literacy

As shown in **Table 3**, compared to people with adequate health literacy, those with limited health literacy received scores with statistically significant lower internal consistency on 5 of the 13 subscales: information avoidance for colon cancer (0.80 vs. 0.88, $p < .001$), unpredictability for diabetes

TABLE 2

Sample Characteristics (*N* = 1,005)

Characteristic	*n*	%
Sex		
Male	484	48
Female	521	52
Age, years		
18-24	86	9
25-34	153	15
35-44	135	13
45-54	175	17
55-64	218	22
65+	238	24
Race/ethnicity		
White, not Hispanic	741	74
Black, not Hispanic	98	10
Hispanic	97	9
Race/ethnicity unkown	69	7
Education		
Less than high school	77	8
High school	295	29
Some college	283	28
Bachelor's degree or higher	350	35
Health literacy		
Limited	206	20
Adequate	799	80

(0.84 vs. 0.88, $p = .01$), perceived severity for diabetes (0.66 vs. 0.75, $p = .003$), need for cognition (0.63 vs. 0.82, $p < .001$), and social desirability (0.52 vs. 0.68, $p < .001$). In contrast, compared to people with adequate health literacy, those with limited health literacy had higher internal consistency for present orientation (0.69 vs. 0.61, $p = .01$) and locus of control (0.82 vs. 0.77, $p = .03$). There were no significant differences in internal consistency for any of the other scales.

Education

Compared to people with education above high school, those who only completed high school or less yielded significantly lower internal consistency on 4 of the 13 subscales: perceived severity for diabetes (0.70 vs. 0.75, $p = .03$), present orientation (0.60 vs. 0.66, $p = .04$), need for cognition (0.73 vs. 0.80, $p < .001$), and social desirability (0.61 vs. 0.70, $p = .04$). In contrast, internal consistency was greater for those who completed high school or less than for those with more than a high school education on 2 of the 13 subscales scales: information avoidance for diabetes (0.86 vs. 0.83, $p = .04$) and unpredictability for colon cancer (0.90 vs. 0.86, $p = .001$). There were no other significant differences in internal consistency for any of the other scales.

DISCUSSION

Our hypothesis received mixed support across three key findings. First, consistent with our hypothesis, internal consistency for three subscales was significantly lower among people with lower education and/or limited health literacy compared to people with higher education and adequate health literacy. Specifically, the need for cognition and perceived severity of diabetes subscales had poor internal consistency among people with limited health literacy. The internal consistency for the social desirability and the two time-orientation subscales were relatively low across all groups compared to other scales.

Second, despite a small number of statistically significant differences between groups, internal consistency remained relatively high regardless of health literacy and educational attainment for eight of the subscales: information avoidance (diabetes and colon cancer), cognitive causation (diabetes and colon cancer), unpredictability (diabetes and colon cancer), perceived severity (colon cancer), and internal health locus of control. Shepperd et al. (2016) also found significant differences in the internal consistency of the need for cognition scale between people with more and less formal education. Our findings for the other scales included in this study are novel.

Possible Explanations for the Results

High readability levels might be one explanation for our results. However, exploratory examination of the FKRL of the social desirability, need for cognition, and perceived severity of diabetes scales all indicated that someone who could read at the sixth-grade level should be able to understand the items (Stockmeyer, 2009). The FKRL also indicated that the time orientation subscales could be read by someone with a sixth-grade reading level, and its internal consistency was nearly universally unacceptable. Therefore, we conclude that readability at the sixth-grade level did not contribute to the low internal consistency of those instruments.

Another explanation could be that the construct validity of the instruments is stronger for groups that were involved in their development. That is, items measuring social desirability, need for cognition, and perceived severity of diabetes might hold different meaning and/or relevance for people with high versus low education and with adequate versus

TABLE 3

Cronbach's Alpha for the Scales

Instrument (Number of Items)	Entire Sample	Limited HL	Adequate HL	*p*	High School Education or Less	High School Education or More	*p*
Information avoidance diabetes (8)	0.84	0.85	0.85	.44	0.86	0.83	.04
Information avoidance colon cancer (8)	0.86	0.83	0.88	.002	0.86	0.86	.38
Cognitive causation diabetes (5)	0.97	0.96	0.96	.12	0.97	0.96	.06
Cognitive causation colon cancer (5)	0.97	0.97	0.96	.32	0.96	0.97	.06
Unpredictability diabetes (3)	0.87	0.82	0.88	< .001	0.88	0.86	.83
Unpredictability colon cancer (3)	0.87	0.87	0.87	.48	0.90	0.86	.001
Perceived severity diabetes (5)	0.73	0.69	0.75	.04	0.70	0.75	.03
Perceived severity colon cancer (5)	0.73	0.74	0.70	.12	0.71	0.74	.09
Present orientation (3)	0.64	0.70	0.61	.01	0.60	0.66	.04
Future orientation (3)	0.68	0.64	0.69	.10	0.65	0.69	.09
Internal health locus of control (6)	0.78	0.81	0.77	.07	0.80	0.77	.08
Need for cognition (7)	0.80	0.66	0.82	< .001	0.73	0.80	< .001
Social sesirability (8)	0.65	0.53	0.68	.001	0.61	0.70	.04

Note. p values are based on the Feldt test comparing alphas across the two levels of HL or education. HL = health literacy.

limited health literacy. For example, the version of the "need for cognition" scale that we administered was developed using college students majoring in arts, business, or social sciences (Sherrard & Czaja, 1999). The social desirability instrument was also developed using college students (Paulhus, 1991). College students are different from the general population across a wide variety of characteristics and experiences (Hanel & Vione, 2016), which could lead them to think about the items differently from people with no college experience. This may help explain why the need for cognition scale had significantly higher internal consistency among people with high education and adequate health literacy than among those with low education and limited health literacy. Although the internal consistency for the time-orientation scales was unacceptable overall and by health literacy and education, they were acceptable among African Americans (alpha = 0.77 and 0.74 for future and present orientation, respectively). The scale was developed in African American women from urban housing communities with low-income (Lukwago et al., 2001).

The information avoidance, cognitive causation, and unpredictability scales were developed and validated with people from a variety of backgrounds, including undergraduate college students, people with low socioeconomic status (Hay et al., 2014; Howell & Shepperd, 2016), people residing in rural areas (Howell & Shepperd, 2016), and patients from an urban primary care clinic (Hay et al., 2014). Thus, it is less surprising that these scales showed good internal consistency across health literacy and education levels. The perceived severity scale was developed with people from seven illness groups in New Zealand and the United Kingdom (Moss-Morris et al., 2002). The internal health locus of control scale was developed among a sample composed of 90% White people and with 74% having at least some college education (Wallston et al., 1978). Future research should investigate whether unacceptable internal consistency of some scales among people with limited health literacy, education, or both is due to differing levels of meaningfulness and relevance of the items for certain populations. One way to address these problems would be to develop new scales with alternative conceptualizations of the constructs through conducting cognitive interviews, particularly among populations with low education and health literacy.

Although the cognitive causation scales for diabetes and colon cancer were skewed (**Table B**) and associated with health literacy (**Table C**), the alphas for these two scales did not significantly differ by health literacy level. Furthermore, they exhibited high internal consistency among those with adequate and limited health literacy (0.96-0.97). None of the other scales exhibited strong skewness. Thus, floor or ceiling effects are unlikely to have influenced our findings.

STUDY LIMITATIONS

First, although we can say that several instruments had unacceptable internal consistency, either in specific demographic groups or in the overall sample, we cannot draw conclusions about the instruments' overall validity or reliability. Second, several instruments have multiple versions with different numbers of items, and we only tested one version. If an instrument has more items, it tends to have higher internal consistency (Cortina, 1993). Therefore, it remains unknown whether the low internal consistency problem is unique for that version or generalizes to other versions. Third, we used the NVS to measure health literacy; however, there are many other health literacy instruments such as the Test of Functional Health Literacy in Adults (Parkert et al., 1995), the Rapid Estimate of Adult Literacy in Medicine (Haun et al., 2014), the 14-item Health Literacy Scale (Suka et al., 2013), and the European Health Literacy Survey (Sørensen et al., 2015). The results might be slightly different if we assess health literacy using other measures because each measure may assess different skills related to health literacy (Haun et al., 2014). Last, we administered NVS by computer whereas the original NVS was developed as a one-on-one interviewer-administered instrument; however, recent studies indicated that NVS can be administered by computer (Mansfield et al., 2018; Weiss, 2018).

CONCLUSIONS

We compared the internal consistency of eight psychological instruments used in health research in participants with adequate and limited health literacy as well as higher and lower education. We conclude that the social desirability and the need for cognition instruments may not accurately measure their target constructs in groups with limited education or health literacy. These findings demonstrate the need for development of new scales in vulnerable populations.

Researchers should be mindful that scales with acceptable internal consistency in their sample as a whole may have unacceptable internal consistency in some sample subsets. Such undesirable variability in internal consistency could undermine instruments' ability to detect phenomena that do, in fact, exist in nature. Considering the importance of psychological instruments for research and practice in psychology and public health (Sturm & Ash, 2005), and ensuring that measures of psychological constructs have acceptable internal consistency among those with limited formal education and health literacy may increase the applicability of research and practice to those groups and thereby alleviate, or at least prevent the exacerbation of, health disparities (Ramírez et al., 2005).

REFERENCES

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology, 67*(2), 319-333. https://doi.org/10.1037/0022-3514.67.2.319

Cohen, A. R. (1957). Need for cognition and order of communication as determinants of opinion change. In C. I. Hovland (Ed.), *The order of presentation in persuasion* (pp. 79-97). Yale University Press.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *The Journal of Applied Psychology, 78*(1), 98-104. https://doi.org/10.1037/0021-9010.78.1.98

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Harcourt Press

Feldt, L. S. (1969). A test of the hypothesis that Cronbach's alpha or Kuder-Richardson coefficent twenty is the same for two tests. *Psychometrika, 34*(3), 363-373. https://doi.org/10.1007/BF02289364

Feldt, L. S., & Kim, S. (2006). Testing the difference between two alpha coefficients with small samples of subjects and raters. *Educational and Psychological Measurement, 66*(4), 589-600. https://doi.org/10.1177/0013164405282488

Ghaddar, S. F., Valerio, M. A., Garcia, C. M., & Hansen, L. (2012). Adolescent health literacy: The importance of credible sources for online health information. *The Journal of School Health, 82*(1), 28-36. https://doi.org/10.1111/j.1746-1561.2011.00664.x PMID:22142172

Gjersing, L., Caplehorn, J. R., & Clausen, T. (2010). Cross-cultural adaptation of research instruments: Language, setting, time and statistical considerations. *BMC Medical Research Methodology, 10*(1), 13. https://doi.org/10.1186/1471-2288-10-13 PMID:20144247

Griffey, R. T., Melson, A. T., Lin, M. J., Carpenter, C. R., Goodman, M. S., & Kaphingst, K. A. (2014). Does numeracy correlate with measures of health literacy in the emergency department? *Academic Emergency Medicine, 21*(2), 147-153. https://doi.org/10.1111/acem.12310 PMID:24673670

Hanel, P. H., & Vione, K. C. (2016). Do student samples provide an accurate estimate of the general public? *PLoS One, 11*(12), e0168354. https://doi.org/10.1371/journal.pone.0168354 PMID:28002494

Haun, J. N., Valerio, M. A., McCormack, L. A., Sørensen, K., & Paasche-Orlow, M. K. (2014). Health literacy measurement: an inventory and descriptive summary of 51 instruments. *Journal of Health Communication, 19*(suppl. 2), 302-333. https://doi.org/10.1080/10810730.2014.936571 PMID: 25315600

Hay, J. L., Baser, R., Weinstein, N. D., Li, Y., Primavera, L., & Kemeny, M. M. (2014). Examining intuitive risk perceptions for cancer in diverse populations. *Health Risk & Society, 16*(3), 227-242. https://doi.org/10.1080/13698575.2014.911822 PMID:24999304

Henson, R. K., Kogan, L. R., & Vacha-Haase, T. (2001). A reliability generalization study of the teacher efficacy scale and related instruments. *Educational and Psychological Measurement, 61*(3), 404-420. https://doi.org/10.1177/00131640121971284

Henson, R. K., & Thompson, B. (2002). Characterizing measurement error in scores across studies: Some recommendations for conducting "Reliability Generalization" studies. *Measurement & Evaluation in Counseling & Development, 35*(2), 113-127. https://doi.org/10.1080/07481756.2002.12069054

Higgins, E. T., Friedman, R. S., Harlow, R. E., Idson, L. C., Ayduk, O. N., & Taylor, A. (2001). Achievement orientations from subjective histories of success: Promotion pride versus prevention pride. *European Journal of Social Psychology, 31*(1), 3-23. https://doi.org/10.1002/ejsp.27

Howell, J. L., & Shepperd, J. A. (2016). Establishing an information avoidance scale. *Psychological Assessment, 28*(12), 1695-1708. https://doi.org/10.1037/pas0000315 PMID:27099980

Hudon, C., Fortin, M., Poitras, M.-E., & Almirall, J. (2012). The relationship between literacy and multimorbidity in a primary care setting. *BMC Family Practice, 13*(1), 33. https://doi.org/10.1186/1471-2296-13-33 PMID:22536833

Hussey, I., & Hughes, S. (2020). Hidden invalidity among 15 commonly used measures in social and personality psychology. A*dvances in Methods and Practices in Psychological Science, 3*, 166-184. https://doi.org/10.1177/2515245919882903

Kindig, D. A., Panzer, A. M., & Nielsen-Bohlman, L. (Eds). (2004). *Health literacy: A prescription to end confusion*. National Academies Press website: https://www.nap.edu/catalog/10883/health-literacy-a-prescription-to-end-confusion

Kutner, M., Greenberg, E., Jin, Y., & Paulsen, C. (2006). *The health literacy of America's adults: Results from the 2003 National Assessment of Adult Literacy*. (NCES 2006-483). U.S. Department of Education website: https://nces.ed.gov/pubs2006/2006483.pdf

Lukwago, S. N., Kreuter, M. W., Bucholtz, D. C., Holt, C. L., & Clark, E. M. (2001). Development and validation of brief scales to measure collectivism, religiosity, racial pride, and time orientation in urban African American women. *Family & Community Health, 24*(3), 63–71. https://doi.org/10.1097/00003727-200110000-00008 PMID:11563945

Mansfield, E. D., Wahba, R., Gillis, D. E., Weiss, B. D., & L'Abbé, M. (2018). Canadian adaptation of the Newest Vital Sign©, a health literacy assessment tool. *Public Health Nutrition, 21*(11), 2038–2045. https://doi.org/10.1017/S1368980018000253 PMID:29692290

Moss-Morris, R., Weinman, J., Petrie, K., Horne, R., Cameron, L., & Buick, D. (2002). The revised illness perception questionnaire (IPQ-R). *Psychology & Health, 17*(1), 1-16. https://doi.org/10.1080/08870440290001494

Onwuegbuzie, A. J., & Daniel, L. G. (2002). A framework for reporting and interpreting internal consistency reliability estimates. *Measurement & Evaluation in Counseling & Development, 35*(2), 89-103. https://doi.org/10.1080/07481756.2002.12069052

Orom, H., Schofield, E., Kiviniemi, M. T., Waters, E. A., Biddle, C., Chen, X., Yuelin, L., Kaphingst, K.A., & Hay, J. L. (2018). Low health literacy and health information avoidance but not satisficing help explain don't know responses to questions assessing perceived risk. *Medical Decision Making, 38*(8), 1006-1017. https://doi.org/10.1177/0272989X18799999 PMID:30403579

Parker, R. M., Baker, D. W., Williams, M. V., & Nurss, J. R. (1995). The test of functional health literacy in adults: A new instrument for measuring patients' literacy skills. *Journal of General Internal Medicine, 10*(10), 537-541. https://doi.org/10.1007/BF02640361 PMID:8576769

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of Social Psychological Attitudes:* Vol. 1. Measures of personality and social psychological attitudes (pp. 17-59). Academic Press. https://doi.org/10.1016/B978-0-12-590241-0.50006-X

Protheroe, J., Whittle, R., Bartlam, B., Estacio, E. V., Clark, L., & Kurth, J. (2017). Health literacy, associated lifestyle and demographic factors in adult population of an English city: A cross-sectional survey. *Health Expectations, 20*(1), 112-119. https://doi.org/10.1111/hex.12440 PMID:26774107

Ramírez, M., Ford, M. E., Stewart, A. L., & A. Teresi, J. (2005). Measurement issues in health disparities research. *Health Services Research, 40*(5 part 2), 1640-1657. https://doi.org/10.1111/j.1475-6773.2005.00450.x PMID:16179000

Reinhardt, B. M. (1991). Factors affecting coefficient alpha: A mini Monte Carlo study. In B. Thompson (Ed.), *Advances in social science methodology* (pp. 3-20). JAI Press.

Shealy, K. M., & Threatt, T. B. (2016). Utilization of the Newest Vital Sign (NVS) in practice in the United States. *Health Communication, 31*(6), 679-687. https://doi.org/10.1080/10410236.2014.990079 PMID:26507669

Shepperd, J. A., Emanuel, A. S., Dodd, V. J., & Logan, H. L. (2016). The reliability of psychological instruments in community samples: A cautionary note. *Journal of Health Psychology, 21*(9), 2033-2041. https://doi.org/10.1177/1359105315569859 PMID:27535831

Sherrard, M., & Czaja, R. (1999). Extending two cognitive processing scales need for cognition and need for evaluation for use in a health intervention. In B. Dubois, T. M. Lowrey, & J. Shrum (Eds.), *European advances in consumer research* (pp. 135-142). Association for Consumer Research. https://www.acrwebsite.org/volumes/11129/volumes/e04/E-04

Sørensen, K., Pelikan, J. M., Röthlin, F., Ganahl, K., Slonska, Z., Doyle, G., Fullam, J., Kondilis, B., Agrafiotis, D., Uiters, E., Falcon, M., Mensing, M., Tchamov, K., van den Broucke, S., Brand, H., & the HLS-EU Consortium. (2015). Health literacy in Europe: Comparative results of the European health literacy survey (HLS-EU). *European Journal of Public Health, 25*(6), 1053-1058. https://doi.org/10.1093/eurpub/ckv043 PMID:25843827

Stockmeyer, N. O. (2009). Using Microsoft Word's readability program. *Michigan Bar Journal, 88*, 46-47.

Sturm, T., & Ash, M. G. (2005). Roles of instruments in psychological research. *History of Psychology, 8*(1), 3-34. https://doi.org/10.1037/1093-4510.8.1.3 PMID:16021762

Suka, M., Odajima, T., Kasai, M., Igarashi, A., Ishikawa, H., Kusama, M., Nakayama, T., Sumitani, M., & Sugimori, H. (2013). The 14-item health literacy scale for Japanese adults (HLS-14). *Environmental Health and Preventive Medicine, 18*(5), 407-415. https://doi.org/10.1007/s12199-013-0340-z PMID:23689952

Taple, B. J., Griffith, J. W., & Wolf, M. S. (2019). Interview administration of PROMIS Depression and Anxiety short forms. *Health Literacy Research and Practice, 3*(3), e196-e204. https://doi.org/10.3928/24748307-20190626-01 PMID:31511845

Taras, V., Rowney, J., & Steel, P. (2009). Half a century of measuring culture: Review of approaches, challenges, and limitations based on the analysis of 121 instruments for quantifying culture. *Journal of International Management, 15*(4), 357-373. https://doi.org/10.1016/j.intman.2008.08.005

U.S. Department of Health and Human Services. (2008). America's health literacy: *Why we need accessible health information*. https://www.ahrq.gov/sites/default/files/wysiwyg/health-literacy/dhhs-2008-issue-brief.pdf

Wallston, K. A., Wallston, B. S., & DeVellis, R. (1978). Development of the Multidimensional Health Locus of Control (MHLC) scales. *Health Education Monographs, 6*(2), 160-170. https://doi.org/10.1177/109019817800600107 PMID:689890

Waters, E. A., Biddle, C., Kaphingst, K. A., Schofield, E., Kiviniemi, M. T., Orom, H., Yuelin, L., & Hay, J. L. (2018). Examining the interrelations among objective and subjective health literacy and numeracy and their associations with health knowledge. *Journal of General Internal Medicine, 33*(11), 1945-1953. https://doi.org/10.1007/s11606-018-4624-2. PMID:30120636

Weiss, B. D. (2018). The Newest Vital Sign: frequently asked questions. *Health Literacy Research and Practice, 2*(3), e125-e127. https://doi.org/10.3928/24748307-20180530-02 PMID:31294286

Weiss, B. D., Mays, M. Z., Martz, W., Castro, K. M., DeWalt, D. A., Pignone, M. P., Mockbee, J., & Hale, F. A. (2005). Quick assessment of literacy in primary care: The newest vital sign. *Annals of Family Medicine, 3*(6), 514-522. https://doi.org/10.1370/afm.405 PMID:16338915

TABLE A

Cronbach's Alpha Across Four Education Levels

Instrument	Education Level			
	Less than High School	High School	Some College	Bachelor's Degree or Higher
Information avoidance diabetes	0.88	0.85	0.78	0.85
Information avoidance colon cancer	0.84	0.87	0.83	0.87
Cognitive causation diabetes	0.96	0.97	0.96	0.96
Cognitive causation colon cancer	0.97	0.96	0.96	0.97
Unpredictability diabetes	0.87	0.88	0.87	0.85
Unpredictability colon cancer	0.90	0.89	0.89	0.84
Perceived severity diabetes	0.73	0.69	0.72	0.77
Perceived severity colon cancer	0.74	0.70	0.74	0.74
Present orientation	0.62	0.59	0.68	0.61
Future orientation	0.59	0.63	0.62	0.62
Internal health locus of control	0.84	0.78	0.77	0.77
Need for cognition	0.76	0.72	0.77	0.81
Social desirability	0.65	0.61	0.69	0.66

TABLE B

Score Distributions

Instrument	Possible Range	Observed Range	Mean	Median	*SD*
Information avoidance diabetes	1-4	1-4	1.95	2	0.59
Information avoidance colon cancer	1-4	1-4	1.96	2	0.62
Cognitive causation diabetes	1-4	1-4	1.37	1	0.62
Cognitive causation colon cancer	1-4	1-4	1.40	1	0.64
Unpredictability diabetes	1-4	1-4	2.87	3	0.75
Unpredictability colon cancer	1-4	1-4	3.19	3	0.67
Perceived severity diabetes	1-4	1-4	2.97	3	0.58
Perceived severity colon cancer	1-4	1-4	3.32	3.4	0.56
Present orientation	3-12	3-12	5.95	6	1.74
Future orientation	4-16	4-16	10.99	11	2.23
Internal health locus of control	6-24	6-24	17.29	17	2.77
Need for cognition	4-28	7-28	19.93	20	3.60
Social desirability	0-8	0-8	3.11	3	1.85

TABLE C

Correlation Between Newest Vital Sign and Scale Scores

Instrument	*r*	*p*
Information avoidance diabetes	–0.13	< .001
Information avoidance colon cancer	–0.13	< .001
Cognitive causation diabetes	–0.30	< .001
Cognitive causation colon cancer	–0.28	< .001
Unpredictability diabetes	–0.11	< .001
Unpredictability colon cancer	0.06	.067
Perceived severity diabetes	0.13	< .001
Perceived severity colon cancer	0.29	< .001
Present orientation	–0.25	< .001
Future orientation	0.03	.392
Internal health locus of control	0.02	.466
Need for cognition	0.24	< .001
Social desirability	–0.03	.316

Note. Newest Vital Sign ranges from 0 to 6.