# On the Validity of the Rapid Estimate of Adult Literacy in Medicine (REALM) Scale as a Measure of Health Literacy

**Levent Dumenci**[*], **Robin K. Matsuyama**, **Laura Kuhn**, **Robert A. Perera**, and **Laura A. Siminoff**
Virginia Commonwealth University

## Abstract

Originally developed to measure the literacy level of patients, the Rapid Estimate of Adult Literacy in Medicine (REALM) scale is one of the most widely used instruments to measure the construct of health literacy. This article critically examines the validity of the REALM as a measure of health literacy. Logical analysis of content coverage led to the conclusion that scores on the REALM should not be used to make inferences about a person's level of health literacy. Rather, the REALM should be used to make inferences about the ability of a person to read and pronounce health related terms. Evidence from an analysis of a sample of 1,037 respondents to the REALM with a cancer diagnosis supports the quality of the REALM as a measure of reading and pronunciation ability. Other uses of the REALM in health literacy research are discussed.

Over the past two decades, health literacy has been an important construct used in research in the field of communication. For example, Kalichman (2008) describes a communication intervention study on health literacy in the prevention and treatment of AIDS. Davis, Crouch, and Long et al. (2002) examined the influence of health literacy in patient-physician communication. And research focused on disease-specific communication has examined the role of patients' health literacy on such health outcomes as diabetes (Schillinger, Piette, & Grumbach et al., 2003) and cancer (Williams, Davis, Parker, & Weiss, 2002).

The definition of health literacy offered by the Institute of Medicine (IOM) has received broad acceptance among clinicians and researchers alike: "The degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decision" (Ratzan & Parker, 2000). As the search for a better definition continues (Berkman, Davis, & McCormack, 2010; Frisch, Camerini, Diviani, & Schultz, 2011; Nutbeam, 2008), this definition of health literacy plays a dominant role in health communication research. Systematic reviews consistently identified several correlates of health literacy. For example, health literacy correlates positively with obtaining diagnostic tests (e.g., mammography), receiving preventive care (e.g., vaccination), and overall health and negatively with using emergency services and mortality (DeWalt, Berkman, Sheridan, Lohr, & Baker, 2004; Paasche-Orlow & Wolf, 2007). A recent study provides a comprehensive overview of the linkage between poor health outcomes and lower health care services utilization and low health literacy (Berkman, et al., 2011). Given health literacy's broad correlates of health outcomes, it is currently viewed as a key social determinant of health (Sentell, Baker, Onaka, & Brown, 2011).

[*]Address correspondence to: Department of Social and Behavioral Health Virginia Commonwealth University One Capitol Square, 9th Floor 830 East Main Street P.O. Box 980149 Richmond, VA 23298 ldumenci@vcu.edu Phone (Work): 804 628-7529 Fax: 804 828-5440.

The Test of Functional Health Literacy in Adults (TOFHLA; Parker, Baker, Williams, & Nurss, 1995) and the Rapid Estimate of Adult Literacy in Medicine (REALM; Davis et al., 1991) are the most widely used instruments to measure health literacy (Berkman, Sheridan, Donahue, Halpern, & Crotty, 2011). Over two-thirds of studies examining the correlates of health literacy used either or both instruments (Paasche-Orlow, Parker, Gazmamarian, Nielsen-Bohlman, & Rudd, 2005). The TOFHLA is a reading comprehension test. It also has a subscale designed to measure numeracy skills. Its short version (S-TOFHLA) is widely used in research settings. The REALM consists of 125 medical terms taken from printed patient education materials. The test takers are asked to read the words aloud in the order of increasing difficulty.

In this manuscript, we focus on the REALM, which gained popularity in health literacy research due in large part to the relatively short time (typically under 5 minutes) it requires for administering and scoring. A person's REALM score is simply the number of correctly pronounced words, which can then be converted into five reading levels as grade equivalencies if desired: 3rd grade and below, 4th to 6th grade, 7th to 8th grade, and 9th grade or above. The test authors contended that patients with less than 9th grade level "will probably have difficulty comprehending most patient education materials" (Murphy, Davis, Long, Jackson, & Decker, 1993; p. 126).

To decrease the test administration and scoring time to about two minutes, Davis et al. (1993) reduced the test items (the number of words) from 125 to 66 and labeled it as the shortened REALM. Over the past decade, two studies have independently proposed much shorter versions of the REALM, one by Bass, Wilson, and Griffith (2003) called the Rapid Estimate of Adult Literacy in Medicine-Revised (REALM-R) involving eight items and another by Arozullah et al. (2007) called the REALM-Short Form (REAM-SF) involving seven items. Yet the 66-item shortened REALM remains most widely used version in research settings, and it is the version we examine in this study.

The content coverage of the shortened REALM was derived from sampling words from medical forms and patient educational materials (Davis et al., 1991). Davis et al. (1993) reported a near perfect test-retest correlation ($r = .99$) suggesting that the REALM scores are very stable over a one-week period. Evidence for concurrent validity was obtained by correlating the REALM scores with scores obtained from three standardized reading tests: the reading recognition section of the Peabody Individual Achievement Test-Revised (PIAT-R), the Wide Range Achievement Test-Revised (WRAT-R), and the Slosson Oral Reading Test-Revised (SORT-R). The validity coefficients were quite high ranging from .88 to .97 (Davis et al., 1993).

The purpose of this study is to critically examine the validity of the shortened REALM (to be referred to as REALM thereafter) as a measure of health literacy. Specifically, we examine the content coverage, internal consistency, item-corrected total correlations, unidimensionality of test scores using factor analysis and item difficulty and discrimination parameters from a single-factor 2-parameter item response theory model (2-PL IRT). As will be seen, although the REALM has certain psychometric qualities that are desirable, we nevertheless argue that it fails as a measure of health literacy and recommend that it not be described and used as such.

## METHOD

### Participants

As part of a project to develop an instrument to measure cancer health literacy, data for this study were collected from individuals 18 years of age or older with a cancer diagnosis.

Exclusion criteria included if participants had an Eastern Cooperative Oncology Group (ECOG) score of four or higher, had been referred to hospice, or their oncologist deemed the participant too sick to complete the task.

The REALM was administered to 1,037 individuals scheduled for a clinic visit at an urban cancer center in the Mid-Atlantic region. Participants were 56.1% female, 36.3% African American, 63.0% non-Hispanic White, and 0.7% multi-racial or other race. One person refused to disclose ethnicity. The mean age was 58.1 years ($sd$ = 12.0). Education levels were 11.7% with less than high school diploma, 16.7% GED or high school diploma, 21.0% with some college, 11.2% associate or technical degree, and 39.3% Bachelor's or higher degree. Median income was $50,000. Self-reported cancer types included Hematologic (27.1%), Breast (17.6%), Gynecologic (11.35%), Genitourinary (9.7%), Head and Neck (9.6%), Lung (7.3%), Colorectal (6.1%), Gastrointestinal (4.3%), Endocrine (1.5%), other (1.8%) and 7 (0.7%) patients with unknown primary diagnoses. All stages of cancer were represented with 14.3% reporting stage I, 11.8% stage II, 14.1% stage III, and 16.1% stage IV. The remaining 43.8% reported their stages as either not applicable due to their cancer type or not knowing it. The median number of years since diagnosis was 2.5 (range: 1 – 43) with 4.6% not reporting date of diagnosis. This study was approved by the Virginia Commonwealth University Institutional Review Board.

### Procedure

Participants were recruited largely through medical record reviews and the distribution of flyers advertising the study in various oncology clinics. Permission to contact was obtained from their oncologists for all participants. Eligible patients were mailed a letter inviting them to participate which also provided them with a phone number they could call to opt out. A week following the letter mail date, patients who had not opted out received a follow up phone call in order to provide additional information about the study and to schedule an interview.

Informed consent was obtained and interviews were conducted in private interview rooms. Research staff were trained to administer the REALM according to the administration rules and coached specifically on how to be sensitive to patients appearing to have low literacy. Participants received a laminated copy of the REALM word list. Research staff held the scoring sheet on a clipboard at an angle to prevent distracting the participant by the scoring procedure and read the test instruction verbatim. Participants were compensated $25 for completing the interview which took one hour on average, with the REALM taking approximately five minutes to administer and score.

## RESULTS

We conducted an item analysis of the REALM, assessed its reliability and the fit of a unidimensional factor model, and conducted a logical analysis of the consistency between the widely accepted definition of health literacy provided above and the representativeness of test items in order to assess content validity (Pedhazur & Schmelkin, 1991). Originating from classical test theory (CTT), item-corrected item-total correlations were used to assess the consistency between each test item and the REALM score. The percent-correct item responses were used to estimate the item difficulty parameters from a CTT perspective. Item-level factor analysis was used to test the unidimensionality of the REALM scores using diagonally-weighted least squares estimation method to take into account the binary item distributions. The chi-square test, root mean square error of approximations (RMSEA), comparative fit index (CFI), and Tucker-Lewis index (TLI) were used to assess model fit, with cutoff values proposed by Hu and Bentler (1999) determining good model fit: RMSEA < .06, CFI > .95, and TLI > .95.

A unidimensional two-parameter logistic item response theory (2PL-IRT) model was used to estimate the item difficulty and discrimination parameters for each item. Analogous to the percent-correct item responses in CTT, the item difficulty parameter represents the trait level with a 50-50 chance of correctly answering a question. The item discrimination parameter represents how well the item differentiates individuals along the trait continuum. An item response curve plots the probability of a correct response as a function of trait level. Precision of measurement depends on the trait level. The test information function depicts the level of precision along the trait continuum in IRT models. The CTT and IRT comparisons and a gentle introduction to IRT have been presented by Embretson and Reise (2000) and Reise and Henson (2003). Mplus (version 6.11) was used to estimate the 2-PL and factor models.

## Content Coverage

Understanding printed materials is a building block of health literacy. Comprehension requires reading and understanding of printed text and materials. During the REALM administration, test takers are instructed to read the printed words and then correctly pronounce them. People who cannot read also can neither pronounce nor understand the printed words as both pronunciation and comprehension of printed text require the ability to read. It is in this narrow sense that these two constructs are related. However, the ability to read and correctly pronounce words does not necessarily mean that the test taker understands the meaning of the word. While some people may correctly pronounce words without understanding their meanings, others may not be able to pronounce words yet understand their meanings. Thus, the REALM does not assess the comprehension aspect of health literacy, as the test score does not measure whether or not each word is understood.

Communication is another major health literacy content area in which the REALM has some coverage, with the entire set of test items involving the ability to pronounce printed medical terms. Pronunciation, however, is only one component of oral communication, i.e., speaking. Listening is the other major component of oral communication (Murphy, 1991), which is not covered by the REALM. In terms of the two remaining primary content areas of health literacy, numeracy and information seeking/navigation, the REALM has no item to represent these content areas.

## Classical Test Theory Item Analysis

For the 66 test items, the percent correct item responses and item-corrected total correlations appear in Table 1, columns 1 and 2, respectively. The correct pronunciation rates ranged from 36% to 99% with the median rate of 93%, which suggest that the REALM is primarily comprised of very easy items. The item-corrected total correlations ranged from .28 to .79 with the median of .67. From the classical test theory perspective, high item-corrected total correlations indicate that the test scores discriminate between low and high scoring individuals. The internal consistency of item responses, as quantified by Cronbach's alpha, was .98.

## Factor Analysis

A one-factor model fit the data well: $\chi^2 = 2441.53$, df = 2,079, p < .001; RMSEA = .013, 90% CI = (.011, .015); CFI = .995; TLI = .995. The standardized factor loadings ranged from .67 to .98 with a median of .89, suggesting that the item responses are highly correlated with the underlying trait. Factor loadings are given in Table 1.

### 2-PL IRT Model

Item characteristic curves (ICCs) from the unidimensional 2-PL IRT model appear in Figure 1. With one exception, item difficulties are located below the average trait level and the ICCs have steep slopes suggesting respectively that items are relatively easy to pronounce and that they are strongly related to the underlying trait. The last two columns of Table 1 provide the difficulty and discrimination estimates from the 2-PL IRT model. The median difficulty was 1.64 standard deviations *below* the mean, where the REALM scores provide the most information (or highest precision of measurement) about the trait. The test information function is depicted in Figure 2.

## DISCUSSION

In recent years, there has been an increasing interest in determining correlates, predictors, and outcomes of health literacy, as well as mediators and moderators of the relationships between health literacy and external variables. Progress in these areas will allow for developing effective interventions to alleviate the adverse impact of low health literacy on people's lives. Research efforts, however, may only be fruitful to the extent that standardized instruments used to measure health literacy in fact measure this construct (validity) and do so accurately (reliability). In this study, we critically examined the validity of the Shortened REALM as a measure of health literacy. Statistically, the REALM appears to provide highly reliable data. However, results from the logical analysis of content coverage revealed that the REALM items minimally cover the health literacy domains shared by the IOM and most recent definitions of health literacy (see Berkman, Davis, & McCormack, 2010; Frisch, Camerini, Diviani, & Schultz, 2011; Nutbeam, 2008). Specifically, the test provides no coverage on three primary content areas of health literacy: Comprehension of printed health materials, numeracy, and information seeking/navigation. We recommend that the REALM not be used to measure health literacy.

Our claim that a person's score on the REALM should not be used to make inferences about that person's health literacy may not be surprising given that measuring health literacy was not the intention of test developers. Although the REALM provides reading grade equivalencies, it was not designed to make inferences about grade equivalencies, either. In fact, the test authors specifically warn test administrators that "scores are interpreted as estimates of literacy, not grade equivalence" (Murphy, et al., 1993, p. 126). Along with the content coverage evidence, the results from extensive psychometric analyses do, however, support the validity of inferences about a person's ability to pronounce medical terms. A caveat should be noted: Over 15% of items have answered correctly by 97% or more patients, providing little or no information about their ability to pronounce medical terms.

To emphasize, the lack of construct validity evidence to infer level of health literacy from the REALM scores is not a deficiency of the test per se. Inferring health literacy from the REALM is, in our judgment, simply a misuse of the test scores. This study focuses only on the validity of inferences from the REALM scores as measures of health literacy. Validity of all other inferences from the REALM scores should be carefully investigated in future studies, including the REALM as a measure of literacy, as originally intended by the test authors.

Our conclusion to dismiss the REALM as a measure of health literacy should not be construed as the dismissal of the REALM in health literacy studies. For instance, one may hypothesize a priori that individuals with low levels of health literacy should also have low ability to pronounce printed medical terms. Researchers may then use the REALM to validate measures of health literacy. In fact, treating the ability to read and pronounce printed health related terms as a correlate or predictor of health literacy is a more

scientifically defensible proposition than treating the pronunciation ability of printed health related terms and health literacy as synonyms.

Adequate content coverage is a sine qua non for all standardized instruments. In health literacy research, two new instruments with explicit content coverage should be mentioned. McCormack et al. (2010) recently developed a skills-based health literacy instrument with an explicit domain specification involving print, oral, and internet-based information seeking and their intersection crossed by specific tasks, e.g., identifying and understanding textual health materials, numeracy, and interpreting information from tables, maps, and videos. Hahn, Choi, Griffin, Yost, and Baker (2011) also proposed a new health literacy instrument that uses talking touchscreen technology (Health LiTT) with three specific content areas: prose, document, and quantitative. Future studies are needed to examine the psychometric properties of both instruments in independent samples.

## Acknowledgments

## References

Arozullah AM, Yarnold PR, Bennett CL, Solltysik RC, Wolf MS, Lee SY, Davis TC. Development and validation of a short-form, rapid estimate of adult health literacy in medicine. Medical Care. 2007; 45:1026–1033. [PubMed: 18049342]

Bass BF, Wilson JF, Griffith CH. A shortened instrument for literacy screening. Journal of General Internal Medicine. 2003; 18:1036–1038. [PubMed: 14687263]

Berkman ND, Davis TC, McCormack L. Health literacy: What is it? Journal of Health Communication. 2010; 15:9–19. [PubMed: 20845189]

Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Health literacy interventions and outcomes: An updated systematic review. Annals of Internal Medicine. 2011; 155:97–197. [PubMed: 21768583]

Berkman, ND.; Sheridan, SL.; Donahue, KE.; Halpern, DJ.; Crotty, K.; Holland, A.; Viswanathan, M. Health literacy interventions and outcomes: An updated systematic review. AHRQ; Rockville, MD: 2011.

Davis TC, Crouch MA, Long SW, Jackson RH, Bates P, George RB, Bairnsfather LE. Rapid assessment of literacy levels of adult primary care patients. Family Medicine. 1991; 23:433–435. [PubMed: 1936717]

Davis TC, Long SW, Jackson RH, Mayeaux EJ, George RB, Murphy PW, Crouch MA. Rapid estimate of adult literacy in medicine: a shortened instrument. Family Medicine. 1993; 25:391–395. [PubMed: 8349060]

Davis TC, Williams MV, Marin E, Parker RM, Glass J. Health literacy and cancer communication. CA: A Journal for Clinicians. 2002; 52:134–149.

DeWalt DA, Berkman ND, Sheridan S, Lohr KN, Baker DW. Literacy and health outcomes: A systematic review of the literature. Journal of General Internal Medicine. 2004; 19:1228–1239. [PubMed: 15610334]

Embreston, SE.; Reise, SP. Item response theory for psychologists. Erlbaum; Mahwah, NJ: 2000.

Frish AL, Camerini L, Diviani N, Schultz PJ. Defining and measuring health literacy: How can we profit from other literacy domains? Health Promotion International. 2011; 27:117–126. [PubMed: 21724626]

Hahn EA, Choi SW, Griffin WJ, Yost KJ, Baker DW. Health literacy assessment using talking touchscreen technology (Health LiTT): A new item response theory-based measure of health literacy. Journal of Health Communication. 2011; 16(sup3):150–162. [PubMed: 21951249]

Hu LT, Bentler PM. Cutoff criteria for indexes in covariance structure analysis: Conventional criteria and new alternatives. Structural Equation Modeling. 1999; 6:1–55.

Kalichman, SC. Health literacy and AIDS treatment and prevention. In: Edgar, T.; Noar, SM.; Freimuth, VS., editors. Communication perspectives on HIV/AIDS for the 21st century. Routledge; New York, NY: 2008. p. 329-350.

McCormack L, Bann C, Squiers L, Berkman ND, Squire C, Schillinger D, Hibbard J. Measuring health literacy: A pilot study of a new skill-based instrument. Journal of Health Communication. 2010; 15(S2):51–71. [PubMed: 20845193]

Murphy JM. Oral communication in TESOL: Integrating speaking, listening, and pronunciation. TESOL Quarterly. 1991; 25:51–75.

Murphy PW, Davis TC, Long SW, Jackson RH, Decker BC. Rapid Estimate of Adult Literacy in Medicine (REALM): A quick reading test for patients. Journal of Reading. 1993; 37:124–130.

Nutbeam D. The evolving concept of health literacy. Social Science and Medicine. 2008; 67:2072–2078. [PubMed: 18952344]

Paasche-Orlow MK, Parker RM, Gazmamarian JA, Nielsen-Bohlman LD, Rudd RR. The prevalence of limited health literacy. Journal of General Internal Medicine. 2005; 20:175–184. [PubMed: 15836552]

Paasche-Orlow MK, Wolf MS. The causal pathways linking health literacy to health outcomes. American Journal of Health Behavior. 2007; 31:S19–S26. [PubMed: 17931132]

Parker RM, Baker DW, Williams MV, Nurss JR. The test of functional health literacy in adults: A new instrument measuring patients' literacy skills. Journal of General Internal Medicine. 1995; 10:537–541. [PubMed: 8576769]

Pedhazur, EJ.; Schmelkin, LP. Measurement, design, and analysis: An integrated approach (Student Edition). Lawrence Erlbaum Associates; Hillsdale, NJ: 1991.

Ratzan, SC.; Parker, RM. Introduction. In: Selden, CR.; Zorn, M.; Ratzan, S,C.; Parker, RM., editors. National library of medicine current bibliographies in medicine: Health literacy. National Institutes of Health; U.S. Department of Health and Human Services; Bethesda, MD: 2000. NLM Pub. No. CBM 2000-1

Reise S, Henson JA. A discussion of modern versus traditional psychometrics as applied to personality assessment scales. Journal of Personality Assessment. 2003; 81:93–103. [PubMed: 12946916]

Schillinger D, Piette J, Grumbach K, Wang F, Wilson C, Daher C, Bindman AB. Physician communication with diabetic patients who have low health literacy. Archives of Internal Medicine. 2003; 163:83–90. [PubMed: 12523921]

Sentell T, Baker KK, Onaka A, Braun K. Low health literacy and poor health status in Asian Americans and Pacific Islanders in Hawai'i. Journal of Health Communication. 2011; 16:279–294. [PubMed: 21951258]

Williams MV, Davis T, Parker RM, Weiss B. The role of health literacy in patient-physician communication. Family Medicine. 2002; 34:383–389. [PubMed: 12038721]
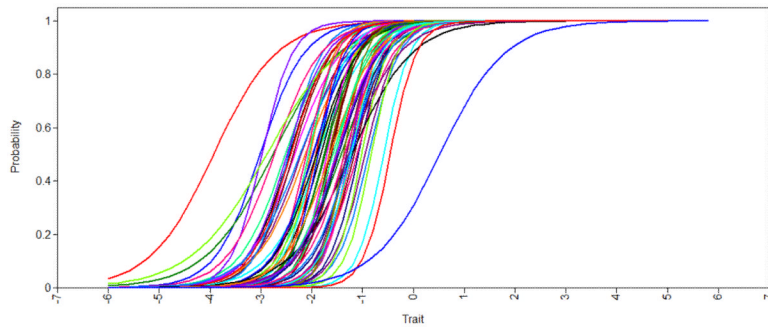
**Figure 1.**
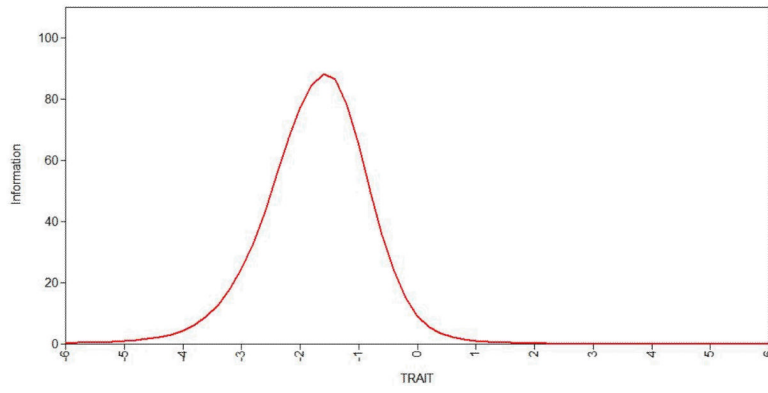Item Characteristic Curves for the 66 Items on the Shortened REALM

**Figure 2.**
Test Information Function

**Table 1**

REALM Item Analysis Statistics

| Item | | % Correct | Item-Corrected Total *r* | Factor Loadings | Difficulty | Discrimination |
|---|---|---|---|---|---|---|
| fat | 1 | .99 | 0.27 | 0.72 | −3.93 | 0.95 |
| flu | 2 | .98 | 0.45 | 0.87 | −3.02 | 1.35 |
| pill | 3 | .96 | 0.41 | 0.70 | −2.77 | 0.90 |
| dose | 4 | .97 | 0.57 | 0.89 | −2.44 | 1.45 |
| eye | 5 | .96 | 0.57 | 0.86 | −2.34 | 1.28 |
| stress | 6 | .97 | 0.59 | 0.90 | −2.36 | 1.55 |
| smear | 7 | .93 | 0.75 | 0.95 | −1.66 | 2.13 |
| nerves | 8 | .95 | 0.64 | 0.88 | −1.97 | 1.51 |
| germs | 9 | .95 | 0.57 | 0.84 | −2.20 | 1.27 |
| meals | 10 | .94 | 0.57 | 0.82 | −2.12 | 1.21 |
| disease | 11 | .97 | 0.56 | 0.88 | −2.52 | 1.31 |
| cancer | 12 | .99 | 0.47 | 0.96 | −2.98 | 1.96 |
| caffeine | 13 | .97 | 0.59 | 0.94 | −2.46 | 1.78 |
| attack | 14 | .98 | 0.52 | 0.88 | −2.74 | 1.27 |
| kidney | 15 | .96 | 0.38 | 0.67 | −2.89 | 0.79 |
| hormones | 16 | .92 | 0.58 | 0.81 | −1.89 | 1.21 |
| herpes | 17 | .91 | 0.69 | 0.89 | −1.50 | 1.76 |
| seizure | 18 | .86 | 0.68 | 0.88 | −1.20 | 1.76 |
| bowel | 19 | .87 | 0.62 | 0.82 | −1.42 | 1.30 |
| asthma | 20 | .94 | 0.67 | 0.90 | −1.82 | 1.64 |
| rectal | 21 | .88 | 0.74 | 0.92 | −1.23 | 2.25 |
| incest | 22 | .86 | 0.69 | 0.88 | −1.17 | 1.84 |
| fatigue | 23 | .89 | 0.75 | 0.93 | −1.31 | 2.21 |
| pelvic | 24 | .85 | 0.71 | 0.90 | −1.08 | 2.10 |
| jaundice | 25 | .82 | 0.66 | 0.87 | −0.95 | 1.91 |
| infection | 26 | .96 | 0.71 | 0.98 | −2.03 | 2.47 |
| exercise | 27 | .96 | 0.69 | 0.96 | −2.06 | 2.13 |
| behavior | 28 | .97 | 0.56 | 0.90 | −2.48 | 1.50 |
| prescription | 29 | .95 | 0.55 | 0.82 | −2.23 | 1.15 |
| notify | 30 | .96 | 0.67 | 0.95 | −2.10 | 2.00 |
| gallbladder | 31 | .94 | 0.67 | 0.90 | −1.88 | 1.54 |
| calories | 32 | .95 | 0.69 | 0.93 | −1.92 | 1.83 |
| depression | 33 | .95 | 0.73 | 0.96 | −1.92 | 2.25 |
| miscarriage | 34 | .95 | 0.73 | 0.96 | −1.88 | 2.25 |
| pregnancy | 35 | .94 | 0.70 | 0.92 | −1.76 | 1.77 |
| arthritis | 36 | .94 | 0.63 | 0.88 | −1.91 | 1.43 |
| nutrition | 37 | .94 | 0.75 | 0.95 | −1.70 | 2.26 |
| menopause | 38 | .93 | 0.75 | 0.94 | −1.64 | 2.06 |
| appendix | 39 | .89 | 0.75 | 0.93 | −1.27 | 2.37 |

| Item | | % Correct | Item-Corrected Total *r* | Factor Loadings | Difficulty | Discrimination |
|---|---|---|---|---|---|---|
| abnormal | 40 | .87 | 0.68 | 0.87 | −1.27 | 1.66 |
| syphilis | 41 | .89 | 0.74 | 0.92 | −1.32 | 2.05 |
| hemorrhoids | 42 | .88 | 0.73 | 0.92 | −1.23 | 2.19 |
| nausea | 43 | .91 | 0.72 | 0.91 | −1.53 | 1.80 |
| directed | 44 | .89 | 0.74 | 0.92 | −1.26 | 2.27 |
| allergic | 45 | .80 | 0.64 | 0.86 | −0.88 | 1.96 |
| menstrual | 46 | .84 | 0.65 | 0.85 | −1.11 | 1.63 |
| testicle | 47 | .84 | 0.71 | 0.91 | −1.02 | 2.32 |
| colitis | 48 | .73 | 0.61 | 0.87 | −0.59 | 2.29 |
| emergency | 49 | .97 | 0.58 | 0.91 | −2.42 | 1.49 |
| medication | 50 | .95 | 0.64 | 0.88 | −1.98 | 1.48 |
| occupation | 51 | .93 | 0.68 | 0.89 | −1.74 | 1.69 |
| sexually | 52 | .81 | 0.51 | 0.72 | −1.21 | 0.99 |
| alcoholism | 53 | .90 | 0.65 | 0.86 | −1.51 | 1.45 |
| irritation | 54 | .93 | 0.76 | 0.95 | −1.61 | 2.19 |
| constipation | 55 | .91 | 0.78 | 0.96 | −1.37 | 2.75 |
| gonorrhea | 56 | .89 | 0.66 | 0.87 | −1.47 | 1.46 |
| inflammatory | 57 | .84 | 0.58 | 0.79 | −1.28 | 1.17 |
| diabetes | 58 | .90 | 0.60 | 0.82 | −1.64 | 1.31 |
| hepatitis | 59 | .91 | 0.66 | 0.86 | −1.58 | 1.54 |
| antibiotics | 60 | .89 | 0.66 | 0.86 | −1.51 | 1.43 |
| diagnosis | 61 | .86 | 0.65 | 0.85 | −1.26 | 1.54 |
| potassium | 62 | .88 | 0.66 | 0.86 | −1.35 | 1.55 |
| anemia | 63 | .80 | 0.68 | 0.90 | −0.83 | 2.33 |
| obesity | 64 | .88 | 0.72 | 0.90 | −1.23 | 1.96 |
| osteoporosis | 65 | .68 | 0.58 | 0.87 | −0.46 | 2.30 |
| impetigo | 66 | .36 | 0.31 | 0.68 | 0.52 | 0.92 |