



Published in final edited form as:

Am J Med Qual. 2006 ; 21(1): 49–56.

Variation in the Readability of Items Within Surveys

José L. Calderón, MD

An assistant professor at the RCMI-Health Services Research Center, Charles R. Drew University of Medicine & Science, Los Angeles, California.

Leo S. Morales, MD, PhD

An associate professor at the David Geffen School of Medicine and School of Public Health, University of California, Los Angeles.

Honghu Liu, PhD and Ron D. Hays, PhD

A professor in the Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles.

Abstract

The objective of this study was to estimate the variation in the readability of survey items within 2 widely used health-related quality-of-life surveys: the National Eye Institute Visual Functioning Questionnaire–25 (VFQ-25) and the Short Form Health Survey, version 2 (SF-36v2). Flesch-Kincaid and Flesch Reading Ease formulas were used to estimate readability. Individual survey item scores and descriptive statistics for each survey were calculated. Variation of individual item scores from the mean survey score was graphically depicted for each survey. The mean reading grade level and reading ease estimates for the VFQ-25 and SF-36v2 were 7.8 (fairly easy) and 6.4 (easy), respectively. Both surveys had notable variation in item readability; individual item readability scores ranged from 3.7 to 12.0 (very easy to difficult) for the VFQ-25 and 2.2 to 12.0 (very easy to difficult) for the SF-36v2. Because survey respondents may not comprehend items with readability scores that exceed their reading ability, estimating the readability of each survey item is an important component of evaluating survey readability. Standards for measuring the readability of surveys are needed.

Keywords

readability; survey research; vulnerable populations

Readability is the semantic and syntactic attributes of the written word. It determines the relative utility of a passage of text for persons with varying degrees of reading skills. Readability assessment has been used for decades to help develop schoolbooks appropriate to grade levels.¹ Readability can be estimated using one of many manual and computerized readability formulas that are based on the number of syllables per word and number of words per sentence a passage contains. These formulae provide an indication of the reading skills needed to decipher and comprehend text.² Word difficulty and sentence length have been found

Corresponding author: José L. Calderón, MD, RCMI-Health Services Research Center, Charles R. Drew University of Medicine & Science, Los Angeles, CA 90059 (e-mail: drcalderon@sbcglobal.net)..

This project was supported by the National Center for Research Resources, Research Centers in Minority Institutions (G12-RR03026-15), UCLA/DREW Project EXPORT, National Institutes of Health, National Center on Minority Health & Health Disparities (P20-MD00148-01), and the UCLA Center for Health Improvement in Minority Elders/Resource Centers for Minority Aging Research, National Institutes of Health, National Institute of Aging (AG-02-004). Leo S. Morales also received partial support from a Robert Wood Johnson Minority Medical Faculty Development Fellowship.

The authors have no affiliation with or financial interest in any product mentioned in this article. The authors' research was not supported by any commercial or corporate entity.

to be the best predictors of the readability of text.³ Text readability is, however, only one of many known factors that facilitate reading. Other factors include the reading skills and motivation of the reader, cultural and social appropriateness, linguistics, formatting, and text legibility.⁴⁻⁶ Readability formulas that use sentence (syntactic) and word (semantic) variables are the most widely used, given their ease of calculation.⁷

Empirical studies of readability measurement date back to the 1920s, when changing social conditions prompted the development of measures of the comprehension of the written language. Increasingly, students from families with limited educational attainment were attending high school. Educators grew concerned that this new class of student was less inclined to use and comprehend vocabulary used in the textbooks of the period and common to families with higher educational attainment.¹ This resulted in the development of several readability formulas. Interestingly, as the social condition of our society again changed after the 1960s, the readability of health communication for primary and secondary prevention of disease among diverse populations became increasingly important.

Interest in the readability of health information emerged in the 1980s. Since then, disparities between the readability of health-related information and the reading skills of patients have been widely reported.⁸⁻¹⁸ Most of these reports are related to the readability of research consent forms and patient education materials. More recently, the readability of Internet-based health information has been reported.^{19,20} Less has been reported on the assessment of survey readability and its importance in survey research.

A literature search on survey readability conducted in May 2005 of Medline (1966-April 2003), CINAHL (1982-April 2003), ClinPSYC (1993-2003), and PsychInfo (2003-2005) databases using keywords *surveys*, *questionnaires*, *readability*, and *comprehension* yielded 17 articles reporting survey readability as one of the study objectives, the formula used for estimating readability, and the readability scores.²¹⁻³⁷ The Flesch-Kincaid (F-K) and the Flesch Reading Ease (FRE) were the formulas most commonly used to assess survey readability (7/17 and 6/17, respectively) (Table 1). However, methods for applying readability formulas to survey text were reported in only 4 of 17 articles (24%), and computerized assessment of readability was reportedly used in 3 of 17 (18%) of the studies. One study reported evaluating the readability of each item separately. None reported variation in the readability of items within surveys.

The average American has a seventh- to eighth-grade reading ability.³⁸ Vulnerable populations (eg, those living in poverty, racial and ethnic minorities, the homeless, persons older than 65 years) are over-represented among those having marginal or very limited reading skills. Knowing survey readability may allow researchers to predict validity of responses to self-administered surveys based on a respondent's reading ability. This notwithstanding, reporting only the overall (mean) readability of a survey may mislead researchers about its usefulness among persons with limited reading skills. Such scores do not reflect the variation in item readability scores within a survey. Unlike prose, in which context can help a reader comprehend the meaning of a passage or sentence, in a survey, the reader is expected to comprehend every item independently. Thus, it is likely to be difficult for the average respondent in the United States to comprehend and respond to surveys that include items that require more than seventh- to eighth-grade reading skills. Surveys with difficult items may be prone to survey nonresponse, item nonresponse, or unreliable responses because of a mismatch between item readability and the reading skills of the respondent.

The purpose of this article is to estimate variation in the readability of survey items within 2 widely used health-related quality-of-life (HRQOL) surveys: the NEI Visual Functioning Questionnaire–25 (VFQ-25) and the Short Form Health Survey, version 2 (SF-36v2). The

VFQ-25 is a disease-targeted instrument designed to assess the influence of eye diseases on HRQOL. The SF-36 is the most widely used generic measure of HRQOL.

METHODS

The VFQ-25 is composed of 28 items, with 1 item being the first general HRQOL item used in the SF-36. Twenty-five of 28 items that comprise the VFQ-25 items are used to derive a VFQ-25 score. They assess 11 vision-related concepts: general vision (1 item), ocular pain (2 items), near vision (3 items), distance vision (3 items), vision-specific social functioning (2 items), vision-specific mental health (4 items), vision-specific role functioning (2 items), dependency due to vision (3 items), driving (3 items), peripheral vision (1 item), and color vision (1 item).³⁹ There are no published readability assessments of the VFQ-25.

The SF-36 is composed of 36 items (1 of which is not used to generate scale scores) addressing 8 health concepts: physical functioning (10 items), bodily pain (2 items), role limitations due to physical health problems (4 items), role limitations due to personal or emotional problems (3 items), emotional well-being (5 items), social functioning (2 items), energy/fatigue (4 items), and general health perceptions (5 items).⁴⁰ Although the readability of version 1 of the SF-36 has been reported to be at the seventh-grade level, an assessment of the variation in the readability across its 36 items has not been reported.

Readability Assessment

The F-K and the FRE formulas were used to estimate the readability of items within the SF-36v2 and the VFQ-25 surveys. The F-K formula rates text on a US grade-school level such that the average eighth grader would be able to read a document that scores 8.0. Scores generated by the F-K formula are highly correlated with scores from other commonly used readability formulas.⁴¹ However, a limitation of the F-K formula in estimating readability is that it has a measurement ceiling of grade level 12. As a result, we also generated FRE scores. The FRE formula rates text on a 100-point scale: the higher the score, the easier it is to understand the document. Both formulas generate scores based on the average number of syllables per word and the number of words per sentence. The correspondence between the scores for these 2 methods and the reading difficulty rating for the scores are displayed in Table 2. Because the readability estimate for a passage is equivalent to the average of the readability of its component sentences, we used the F-K and FRE formulas to assess the readability of single items as well as the survey as a whole.

The formulas used for calculating F-K and FRE scores are F-K reading grade level score $(0.39 \times \text{ASL}) + (11.8 \times \text{ASW}) - 15.59$ and FRE score $206.835 - (1.015 \times \text{ASL}) - (84.6 \times \text{ASW})$, where ASL is the average sentence length (the number of words divided by the number of sentences) and ASW is the average number of syllables per word (the number of syllables divided by the number of words).

We selected these formulas because they are available in Microsoft Word programs and are therefore readily available to nearly all investigators interested in assessing the readability of text. Moreover, the use of computerized software reduces the amount of work required to produce readability estimates, eliminates human error inherent in manual calculation, and requires little training.⁴²

Several items within the surveys presented a unique methodological challenge for assessing readability because of their fragmented formatting. Sentence fragments such as preamble phrases and subsequent question choices that comprise stem-leaf formats cannot be tested for readability using computerized methods. These fragments are not scored because they are not recognized by the computer as a complete statement, declaration, or question that ends with a

period, exclamation mark, or question mark. To overcome this obstacle, we combined preamble phrases with question choices for items using the stem-leaf format to form complete questions. This approach has been used successfully with the FRE and is in keeping with recommendations to test only running text within pieces of writing.^{43,44} Response options tend to be single words or very short phrases and are likely to score as very easy to read; therefore, their readability was not assessed. The input for the readability assessment (raw data) for both surveys is available at <http://www.chime.ucla.edu/measurement/measurement.htm>.

Statistical Analyses

We evaluated the distribution of the readability scores and calculated the mean, standard deviation, median, and range across items. The deviation of item readability estimates from the mean score was examined and displayed graphically. As noted above, the FK formula has a reading grade level measurement ceiling of 12th grade. Items scoring 12 on the F-K formula may actually have a readability score greater than 12th grade; thus, there is potential for underestimating means, standard deviations, medians, and ranges. However, this computerized method still serves as a useful tool for testing readability and guiding the development of information with enhanced readability.

RESULTS

VFQ-25

The mean and median F-K reading grade level scores for the VFQ-25 items were 7.8 and 8.3, classifying them as fairly easy and average to read, respectively. The standard deviation was 2.6, and the range was 3.7 to 12.0 (Table 3). Fourteen of 25 VFQ-25 items had an F-K score above the mean score, and 11 items had an F-K score below the mean score. Eight items scored above 9.0, and 2 items scored 12.0 or greater (Figure 1). This means that 40% (10/25) of the VFQ-25 items may not be readable to persons having less than high school reading skills. The 2 items scoring 12.0 using the F-K formula may actually require college-level reading skills to be comprehended given that the F-K has a ceiling of 12.0. However, the FRE formula does not depend on grade level scores and is a good adjunct to the F-K formula. For the VFQ-25 items, the mean and median FRE scores were 68.4 and 67.3, respectively. The standard deviation was 11.2, and the range was 41.8 to 87.9 (difficult to very easy to read). Given the skewed distribution of item readability in both surveys, the median is a better estimate of the central tendency than the mean.

SF-36v2

The mean and median F-K scores for the SF-36v2 were 6.4 and 4.8, consistent with documents that are easy and very easy to read, respectively. The standard deviation was 3.1, and the range was 2.2 to 12.0: very easy to difficult to read (Table 3). Fifteen of the SF-36 items had an F-K score above the mean score, and 21 were below the mean score. Eight items scored above 9.0, and 5 scored 12.0 (Figure 2). This means that 36% (13/36) of the SF-36v2 items may not be comprehensible by persons having less than high school reading skills. The 5 items scoring 12 may require college-level educational attainment to be read with comprehension given that the F-K formula does not measure readability above 12.0. The mean and median FRE scores for SF-36v2 questions were 76.0 and 75, respectively, with a standard deviation of 19.2 and a range of 31.7 to 100.0.

DISCUSSION

The results of the study reveal notable variation in the readability scores of items in both the VFQ-25 and the SF-36v2. Where the average readability scores for the VFQ-25 and SF-36v2 suggest that these surveys may be appropriate for people with high school-level reading skills,

the variation in the item readability scores suggests that they are not, with some items in each survey requiring college-level reading skills. The readability of the VFQ-25 has not been previously reported. These findings demonstrate that simply reporting survey mean readability does not reflect variation in the readability of individual items. Since respondents are asked to respond to survey items one at a time, the actual level of readability of each item is very important because that is what respondents actually encounter in a survey. Respondents may understand some items but may choose to skip or give invalid responses to items exceeding their reading skills.

Although this study demonstrated notable variation in the readability within 2 widely used HRQOL surveys, further studies are needed to assess the prevalence of this potential problem in other widely used surveys. One implication of our findings is the existence of a survey readability disparity for vulnerable populations who tend to have limited reading skills (eg, the elderly, populations living in poverty, racial and ethnic minorities). Importantly, this applies especially to minority populations who experience health disparities and among whom these surveys may be used to gain a greater understanding of their poor health status.

The recommended readability for documents intended for use among vulnerable populations is an FK score of 5.0 or less or an FRE score of 80 or greater.^{45,46} Indeed, when health care delivery systems and providers prepare written health information for vulnerable populations, it should be at the level of children's books.⁴⁷ Figures 1 and 2 demonstrate that 32% (9/28) of VFQ-25 and 50% (18/36) of SF-36v2 items meet these criteria as measured by the F-K score of 5.0 or lower. Consequently, 19 VFQ-25 and 18 SF-36v2 items may not be comprehensible to respondents with limited literacy skills. This has important implications for the reliability and validity of data collected using these surveys in vulnerable populations who bear the greatest burden of chronic disease and account for the majority of health care expenditures.

Readability formulas can help to achieve writing that is easier to read by estimating text reading difficulty and helping authors to tailor their materials to audiences with varying literacy skills. Importantly, readability formulas should be used as estimates of text readability and not as guides to writing itself. The latter is the domain of what has been termed *wriability*, which is concerned with writing, rewriting, and editing to produce written text that is linguistically appropriate and comprehensible.^{6,7} For example, the words in a paragraph may be scrambled to form a new paragraph that may be incomprehensible, but both paragraphs will have the same readability score. Furthermore, mechanically applying readability formulas to text may lead to poor-quality writing (ie, substituting easier words for harder words and shorter sentences for longer sentences to achieve lower [assumed to be "better"] readability scores). Applying the principles of simplifying text, including the use of commonly used words that tend to be monosyllabic and bisyllabic and using shorter sentences that are simple sentences (1 clause or idea), may help to avoid this limitation.⁴⁸⁻⁵¹

The methods described in this study may be useful for estimating survey readability and for providing guidance to modify existing surveys and develop new surveys so that they may better match the reading skills of respondents. A limitation of this study is that only 2 surveys were examined. However, they have been widely used across populations with varying socioeconomic status and serve as important examples of the need to establish standards for assessing the readability of surveys. Further studies are needed to gauge the extent to which survey questions vary in readability and to assess their impact on data-modifying issues such as bias and item nonresponse.

Because addressing health disparities among vulnerable populations is a national priority, the development and use of HRQOL measures and other surveys that are readable and understandable by these populations are priorities. Using very easy to read text (fifth-grade

reading level) as a standard and assessing reading skills of respondents may open a window of opportunity for enhancing the comprehension of survey items for most populations groups. This also may be a cost-effective approach to survey research because fewer versions of a survey would be needed if the intent were to use the survey cross-culturally and across groups with varying socioeconomic status and reading skills. Given this, we suggest that surveys may not be appropriate for measuring HRQOL for everyone, particularly persons with limited literacy. In the latter case, qualitative methods such as focus groups or in-depth interviews may be more appropriate.

Field testing and qualitatively assessing surveys using focus groups and cognitive interviews add a useful dimension to assessing the linguistic and cultural appropriateness of surveys for specific populations.⁵² Estimating the readability of each survey item is an important component of evaluating surveys, particularly if they are intended for administration cross-culturally and to vulnerable populations that tend to have limited literacy skills. The inclusion of item readability may add an important component in assessing the reliability and validity of a survey, particularly when used among vulnerable populations who tend to experience disparities in health.

REFERENCES

1. Chall, JS. The beginning years. In: Zakaluk, BL.; Samuels, SJ., editors. *Readability, Its Past, Present and Future*. International Reading Association; Newark, Del: 1988. p. 2-13.
2. Smith, S. Readability testing of health information. *Beginnings: a practical guide through pregnancy*. Available at: <http://www.beginningsguides.net/readabil.htm>. Accessed June 12, 2005
3. Meade CD, Smith CF. Readability formulas: cautions and criteria. *Patient Educ Couns* 1991;17:153–158.
4. Pichert JW, Elam P. Readability formulas may mislead you. *Patient Educ Couns* 1985;7:181–191.
5. Boyd MD. A guide to writing effective patient education materials. *Nurs Manage* 1987;18:56–57. [PubMed: 3648599]
6. Doak, CG.; Doak, LG.; Root, JH. *Teaching Patients With Low Literacy Skills*. Lippincott; New York, NY: 1985.
7. Fry EB. Reading formulas: maligned but valid. *Journal of Reading* 1989;32:292–297.
8. Morrow GR. How readable are subject consent forms? *JAMA* 1980;24:56–58. [PubMed: 7382056]
9. Powers RD. Emergency department patient literacy and the readability of patient-directed materials. *Ann Emerg Med* 1988;17:124–126. [PubMed: 3337429]
10. Meade CD, Byrd JC. Patient literacy and the readability of smoking education literature. *Am J Public Health* 1989;79:204–206. [PubMed: 2913844]
11. Tarnowski KJ, Allen DM, Mayhall C, Kelly PA. Readability of pediatric biomedical research informed consent forms. *Pediatrics* 1990;85:58–62. [PubMed: 2296494]
12. Meade CD, Howser DM. Consent forms: how to determine and improve their readability. *Oncol Nurs Forum* 1992;19:1523–1528. [PubMed: 1461766]
13. Hopper KD, Lambe HA, Shirk SJ. Readability of informed consent forms for use with iodinated contrast media. *Radiology* 1993;187:279–283. [PubMed: 8451429]
14. Chacon D, Kisson N, Rich S. Education attainment level of caregivers versus readability level of written instructions in a pediatric emergency department. *Pediatr Emerg Care* 1994;10:144–149. [PubMed: 8058557]
15. Baker LM, Wilson FL. Consumer health materials recommended for public libraries: too tough to read? *Public Libraries* 1996;35:124–30.
16. Ott BB, Hardie TL. Readability of advance directive documents. *Image J Nurs Sch* 1997;29:53–57. [PubMed: 9127541]
17. Baker LM, Wilson FL, Kars M. The readability of medical information on InfoTrac: does it meet the needs of people with low literacy skills? *Reference & User Services Quarterly* 1997;37:155–160.

18. Report of the Council on Scientific Affairs: Health Literacy. *JAMA* 1999;281:552–557. [PubMed: 10022112]
19. Graber MA, Roller CM, Kaeble B. Readability levels of patient education material on the World Wide Web. *J Fam Pract* 1999;48:58–61. [PubMed: 9934385]
20. Berland G, Elliott M, Morales LS, et al. Health information on the Internet: accessibility, quality, and readability in English and Spanish. *JAMA* 2001;285:2612–2621. [PubMed: 11368735]
21. Berndt DF, Schwartz S, Kaiser CF. Readability of self-report depression inventories. *J Consult Clin Psych* 1983;51:627–628.
22. Price JH, O'Connell JK, Kukulka G. Development of a short obesity knowledge scale using four different response formats. *J Sch Health* 1985;55:380–384.
23. Jensen BJ, Witcher DB, Upton LR. Readability assessment of questionnaires frequently used in sex and marital therapy. *J Sex Marital Ther* 1987;13:137–141. [PubMed: 3612824]
24. Devins GM, Binik YM, Mandin H, et al. The Kidney Disease Questionnaire: a test for measuring patient knowledge about end-stage renal disease. *J Clin Epidemiol* 1990;43:297–307. [PubMed: 2313319]
25. Macey BA, Bouman CC. An evaluation of validity, reliability and readability of the Critical Care Family Needs Inventory. *Heart Lung* 1991;20:398–403. [PubMed: 2071431]
26. Heyland DK, Tranmer JE. Measuring family satisfaction with care in the intensive care unit: the development of a questionnaire and preliminary results. *J Crit Care* 2001;16:142–149. [PubMed: 11815899]
27. Paolo AM, Ryan JJ, Dunn GE, Van Fleet J. Reading level of the Dissociative Experiences Scale. *J Clin Psychol* 1993;49:209–211. [PubMed: 8486802]
28. Edlund C. An effective methodology for surveying a Medicaid population: the 1996 Oregon Health Plan Client Satisfaction Survey. *J Ambul Care Manage* 1997;20:37–45. [PubMed: 10164032]
29. MacDiarmid SA, Goodson TC, Holmes TM, Martin PR, Doyle RB. An assessment of the comprehension of the American Uro-logical Association Symptom Index. *Clin Urol* 1997;159:873–874.
30. Beckman HT, Lueger RJ. Readability of self-report clinical outcome measures. *J Clin Psychol* 1997;53:785–789. [PubMed: 9403379]
31. Eaden JA, Abrams K, Mayberry JF. The Crohn's and Colitis Knowledge Score: a test for measuring patient knowledge in inflammatory bowel disease. *Am J Gastroenterol* 1999;94:3560–3566. [PubMed: 10606319]
32. Pande KC, de Takats D, Kanis JA, Edwards V, Slade P, McCloskey EV. Development of a questionnaire (OPQ) to assess patient's knowledge about osteoporosis. *Maturitas* 2000;37:75–81. [PubMed: 11137326]
33. Kimble LP, Dunbar SB, McGuire DB, De A, Fazio S, Strickland O. Cardiac instrument development in a low-literacy population: the revised Chest Discomfort Diary. *Heart Lung* 2001;30:312–320. [PubMed: 11449218]
34. Rowan K. The development and validation of a multi-dimensional measure of chronic foot pain: the Rowan Foot Pain Assessment Questionnaire (ROFPAQ). *Foot Ankle Int* 2001;22:795–809. [PubMed: 11642531]
35. Otley A, Smith C, Nicholas D, et al. The IMPACT Questionnaire: a valid measure of health-related quality of life in pediatric inflammatory bowel disease. *J Pediatr Gastroenterol Nutr* 2002;35:557–563. [PubMed: 12394384]
36. Dolovich LR, Nair KM, Ciliska DK, et al. The Diabetes Continuity of Care Scale: the development and initial evaluation of a questionnaire that measures continuity of care from the patient perspective. *Health Soc Care Community* 2004;12:475–487. [PubMed: 15717895]
37. Travess HC, Newton JT, Sandy JR, Williams AC. The development of a patient-centered measure of the process and outcome of combined orthodontic and orthognathic treatment. *J Orthod* 2004;31:220–234. [PubMed: 15489366]
38. Kirsh, IS.; Jungeblut, A.; Jenkins, L.; Kolstad, A. *Adult Literacy in America: A First Look at the Results of the National Adult Literacy Survey*. US Department of Health, Education and Welfare; Washington, DC: 1993.

39. Mangione CM, Lee PP, Gutierrez PR, Spritzer K, Berry S, Hays RD. Development of the 25-item National Eye Institute Visual Function Questionnaire. *Arch Ophthalmol* 2001;119:1050–1058. [PubMed: 11448327]
40. Hays RD, Sherbourne D, Mazel RM. The RAND 36-item health survey 1.0. *Health Econ* 1993;2:217–227. [PubMed: 8275167]
41. Meade CD, Smith CF. Readability formulas: cautions and criteria. *Patient Educ Couns* 1991;17:153–158.
42. Mailloux SL, Johnson ME, Fisher DG, Pettibone TJ. How reliable is computerized assessment of readability? *Comput Nurs* 1995;13:221–225. [PubMed: 7585304]
43. Flesch, R. How to write plain English. Available at: <http://www.mang.canterbury.ac.nz/courseinfo/AcademicWriting/Flesch.htm>. Accessed October 25, 2003
44. Weiss BD, Coyne C. Communicating with patients who cannot read. *N Engl J Med* 1997;387:272–274. [PubMed: 9227936]
45. Calderon JL, Zadshir A, Norris K. Structure and content of chronic kidney disease information on the World Wide Web: barriers to public understanding of a pandemic. *Nephrol News Issues* 2004;18:76, 78–79. [PubMed: 15551616]81-84
46. Calderón JL, Beltran R. Pitfalls in health communication: health policy, institutions, structure and process of healthcare. *Med General Med* 2004;6:9.
47. Calderón, JL.; Baker, RS.; Hays, RD. Readability of written documents provided to vulnerable patients; Paper presented at: National Meeting of the American Public Health Association; Chicago, Ill. November 2002;
48. Plimpton S, Root J. Materials and strategies that work in low literacy health communication. *Public Health Rep* 1994;109:86–92. [PubMed: 8303020]
49. Doak LG, Doak CC, Meade CD. Strategies to improve cancer education materials. *Oncol Nurs Forum* 1996;23:1305–1312. [PubMed: 8883075]
50. Robinson A, Miller M. Making information accessible: developing plain English discharge instructions. *J Adv Nurs* 1996;24:528–535. [PubMed: 8876413]
51. Weiss BD, Coyne CA. Communicating with patients who cannot read. *N Eng J Med* 1997;337:272–273.
52. Calderón JL, Baker RS, Wolf KE. Focus groups: a qualitative method complementing quantitative research for studying culturally diverse groups. *Educ Health* 2000;13(1):91–95.



Figure 1.
Variation in Visual Functioning Questionnaire–25 readability.

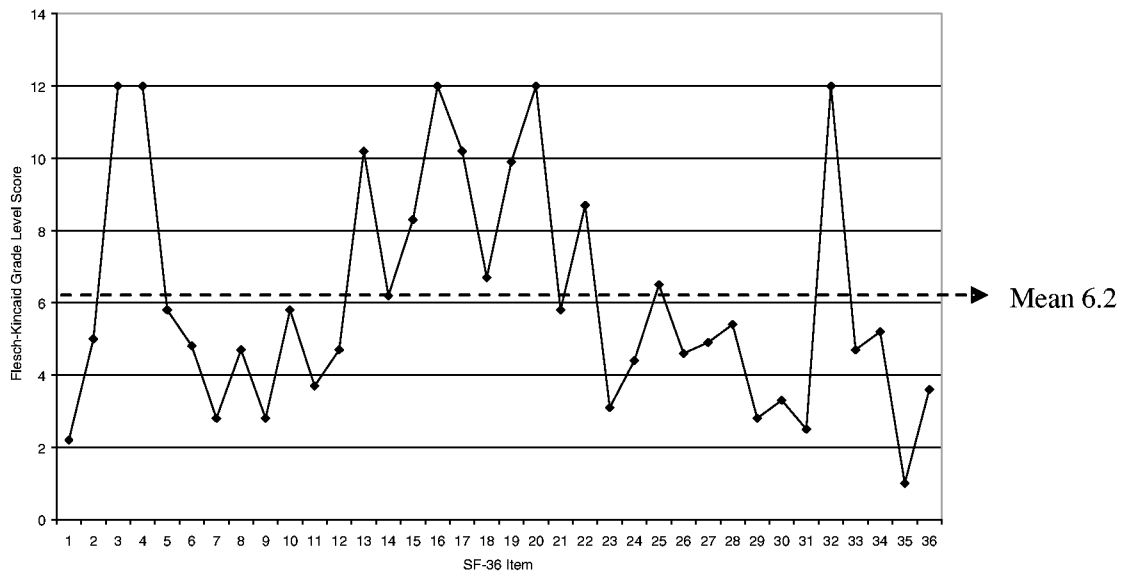


Figure 2.
Variation in Short Form Health Survey item readability.

Table 1
Publications on Survey Readability: Methods, Application to Test, and Scores

Author, Year	Survey Theme	Readability Test	Application	Scores
Berndt, 1983	Depression inventories-5	FRE, Gunning	Not indicated	5th-12th grade
Price, 1985	Obesity	SMOG	Response options	9th grade
Jensen, 1987	Marital surveys-9	Forbes-Cottle	Instructions and questions	6th-college
Devins, 1990	Kidney knowledge	Multiple methods	Not indicated	9th grade
Macey, 1991	Critical care family needs	Gunning-Fox	Not indicated	9th grade
Paolo, 1993	Dissociative experiences	F-K	Questions	10th grade
Beckman, 1997	Clinical outcomes surveys-5	FRE	Instructions and questions	8th-9th grade
Edlund, 1997	Health plan satisfaction	F-K	Not indicated	6th grade
MacDiarmid, 1997	Urological symptoms	Dale-Chall	Survey as a whole	6th grade
Eaden, 1999	Colitis knowledge	FRE	Not indicated	74.3
Pande, 2000	Osteoporosis	FRE	Not indicated	74.3
Heyland, 2001	Intensive care unit satisfaction	F-K	Not indicated	6th grade
Kimble, 2001	Cardiac	F-K	Not indicated	4th grade
Rowan, 2001	Foot pain	FRE	Not indicated	82.4
Otley, 2002	Pediatric inflammatory bowel disease	F-K	Not indicated	5th grade
Dolovich, 2004	Foot pain	FRE	Not indicated	82.4
Travess, 2004	Pediatric inflammatory bowel disease	F-K	Not indicated	5th grade

FRE = Flesch Reading Ease; SMOG = Simple Measure of Gobbledygook; F-K = Flesch-Kincaid Formula.

Table 2
 Reading Difficulty Rating of Flesch Reading Ease Scores and Flesch-Kincaid Grade Level Scores

Reading Difficulty Rating	Flesch Reading Ease Score	Flesch-Kincaid Grade Level Score
Very easy	90-100	5th
Easy	80-90	6th
Fairly easy	70-80	7th
Standard	60-70	8th-9th
Fairly difficult	50-60	10th-12th
Difficult	30-50	13th-16th
Very difficult	0-30	≥College graduate

Table 3
SF-36 and VFQ-25 Mean, Standard Deviation, Median, and Range of Item Readability

	Flesch Reading Ease Index		Flesch-Kincaid Formula	
	SF-36	VFQ-25	SF-36	VFQ-25
Mean	76	68.4	6.2	7.7
Standard deviation	19.2	11.2	3.1	2.7
Median	75	67.3	4.7	8.3
Range	31.7-100	41.8-87.9	2.2-12	3.7-12

SF-36 = Short Form Health Survey; VFQ-25 = Visual Functioning Questionnaire-25.