

2006

How Readable Are Summary Plan Descriptions For Health Care Plans?

Colleen E. Medill

University of Nebraska-Lincoln, cmedill2@unl.edu

Richard L. Weiner

University of Nebraska-Lincoln, rwiener2@unl.edu

Brian H. Bornstein

University of Nebraska-Lincoln, bbornstein2@unl.edu

E. Kiernan McGorty

University of Nebraska-Lincoln

Follow this and additional works at: <http://digitalcommons.unl.edu/lawfacpub>



Part of the [Legal Studies Commons](#)

Medill, Colleen E.; Weiner, Richard L.; Bornstein, Brian H.; and McGorty, E. Kiernan, "How Readable Are Summary Plan Descriptions For Health Care Plans?" (2006). *College of Law, Faculty Publications*. 65.

<http://digitalcommons.unl.edu/lawfacpub/65>

This Article is brought to you for free and open access by the Law, College of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in College of Law, Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

How Readable Are Summary Plan Descriptions For Health Care Plans? p. 2

Executive Summary:

How Readable Are Summary Plan Descriptions For Health Care Plans?

- ***SPDs are the primary source of health plan information:*** The summary plan description (SPD) is the primary source of information for workers who participate in an employment-based health care plan. This study investigates whether private-sector employers' SPDs are written so that an average plan participant can identify and read important information contained in the document, as required by federal law. The study collected and tested SPDs for 40 health care plans from a diverse national sample and subjected them to content and readability analyses.
- ***Important information contained in many SPDs is written at a reading level that may be too high for the average plan participant:*** The study found that the average readability level for important information concerning eligibility, benefits, and participant rights and responsibilities in summary plan descriptions is written at a first year college reading level. The average level of readability for SPDs is higher than the recommended reading level for technical material. Some of the SPDs in the study sample use language written at a 9th grade reading level. Other SPDs use language written at nearly a college graduate (16th grade) reading level.
- ***Readability levels do not differ between single-employer health care plans and multi-employer health care plans:*** The study found no statistically significant differences in readability levels between SPDs for health care plans sponsored by single- and multi-employer health care plans for union workers. The language used for both types of plans tested out at a college reading level.
- ***Fundamental literacy a barrier to health care literacy:*** According to the U.S. Department of Education, about 43 percent of American adults scored at below basic or basic levels of prose literacy (able to read and understand sentences and paragraphs), and 34 percent scored at below basic or basic levels of document literacy (able to read sentence fragments, such as a medical prescription). Given the rapidly rising share of the immigrant, non-English-speaking U.S. population, a major barrier to *health care* literacy is the underlying issue of *fundamental* literacy.
- ***Implications:*** Findings from this study suggest that fundamental improvements are needed in the readability of written SPDs, and that employers and plan administrators should explore the use of alternative methods of communication to plan participants beyond the written SPD. Also, the trend toward consumer-driven health care plans may make the challenge of communicating information to participants through written SPDs even more difficult, since these plans shift significant responsibility to the participants in the plan for decisions concerning the utilization of health care services. This shift in decision-making responsibility to participants makes it more important than ever that participants understand how their health plan works.

■ ***How Readable Are Summary Plan Descriptions For Health Care Plans?***

By Colleen E. Medill, EBRI Fellow, and Richard L. Wiener, Brian H. Bornstein, and E. Kiernan McGorty¹

Introduction

Numerous studies have shown that health care literacy in the United States is poor: A high percentage of Americans do not understand how their health care plans operate (Hibbard & Jewett, 1997; Hibbard et al., 1998; Isaacs, 1996), and also have trouble interpreting medical information, thereby jeopardizing their health, inflating costs, and complicating the work of health professionals (Vastag, 2004). More than 30 percent of English-speaking Americans have inadequate or marginal health literacy as measured by the Test of Functional Health Literacy in Adults or the Short Test of Functional Health Literacy in Adults (Gazmararian et al., 1999; Gazmararian et al., 2003; Williams et al., 1995). Even when controlling for such factors as educational level and socioeconomic status, inadequate health literacy is associated with inferior understanding of health-related information along with worse health, less use of preventive health services, and inflated health care costs (Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs, 1999; Vastag, 2004).

Recently, the issue of health care literacy and the information provided to participants about their health care plans has become much more significant with the introduction of consumer-driven health care plans into the marketplace (Fronstin, 2002 and 2004). Consumer-driven health care plans shift significant responsibility to the participants in the plan for decisions concerning the utilization of health care services (Fronstin and Collins, 2005). This shift in decision-making responsibility to participants makes a consumer-driven health care plan more complex for the participant to navigate, in design and function, than an insured health care plan or a managed health care plan. This shift in responsibility also makes it more important for participants to understand how their health plan works.

The federal Employee Retirement Income Security Act of 1974 (ERISA) requires administrators of private-sector employment-based health care plans to provide important information to plan participants concerning how the plan works. This information is communicated through a written document known as the “summary plan description,” or SPD. The summary plan description is the primary source of information for participants in the health care plan concerning eligibility for health care benefits, the scope of health care benefits covered by the plan, and the rights and responsibilities of participants under the plan.

ERISA requires that the SPD document be written in a manner that is understandable to the average plan participant. The federal regulatory agency charged with overseeing enforcement of this requirement is the Department of Labor. Anecdotal evidence has long suggested that, despite ERISA’s requirement, summary plan descriptions for health care plans often are written in a manner that is difficult for the plans’ participants to read and comprehend. In 2005, the Department of Labor’s Employee Benefits Security Administration formed the Working Group on Health and Welfare Benefit Plans’ Communications to investigate whether summary plan descriptions for health care plans today function as an effective communication tool. The Working Group solicited testimony from experts in the employee benefits community on this topic. Based on this testimony, the Working Group recommended that additional regulatory guidance was needed to help plan administrators prepare readable summary plan descriptions. The Working Group further recommended that new or enhanced regulatory mechanisms were needed to enforce ERISA’s requirement that summary plan descriptions be written in a manner that is understandable to the average plan participant (U.S. Department of Labor, 2005).

Considering that one-third or more of the adult population in the United States has limited ability to read English at all, the inability of large numbers of Americans to understand an SPD should not be surprising. According to the most recent National Assessment of Adult Literacy (NAAL), conducted by the U.S. Department of Education in 2003, about 43 percent of American adults scored at below basic or basic levels of prose literacy (able to read and understand sentences and paragraphs), and 34 percent scored at or below basic levels of document literacy (able to read sentence fragments, such as a medical prescription).²

In addition, the NAAL found that illiteracy levels were highest among those who did not graduate from high school (a group that comprises 15 percent of the adult population), spoke no English before starting school (comprising 13 percent of the adult population), or were Hispanic (12 percent of the adult population). Given the rapidly rising share of the immigrant, non-English-speaking U.S. population (especially among Spanish-speakers), a major and obvious barrier to *health care* literacy is the underlying issue of *fundamental* literacy—defined by the Department of Education as “using printed and written information to function in society.” Due to the inherent complexity of the medical and insurance bureaucracies, the technical and sometimes confusing terminologies they use, and the inability of large numbers of Americans to understand written English at even a basic level or follow written directions, SPD reader comprehension problems are inevitable.

Nevertheless, the readability of SPDs has only recently begun to receive attention. This study provides quantifiable evidence about whether summary plan descriptions cover the necessary topics that participants need to know about their health plans, and the educational level required to read them.

This study finds that in many cases crucial information is written at a reading level that is likely too high for the average plan participant. Eligibility, benefits, and participant rights and responsibilities information in SPDs is written, on average, at a first-year college reading level. The average readability for SPDs is higher than the recommended reading level for technical material. Some of the SPDs in the study sample use language written at a junior high school level (9th grade), but other SPDs use language written at nearly a college graduate (16th grade) level.

These results suggest that fundamental improvements are needed in the readability of written SPDs, and that employers and plan administrators should explore the use of alternative methods of communication to plan participants beyond the written SPD.

Study Methodology

Overview of the Study—This study used empirical techniques from the field of educational psychology to investigate the likelihood that an average plan participant could identify and read important information contained in a summary plan description for a health care plan. The study consisted of two stages:

- The first stage of the study used experts to identify important information in the text of the summary plan description document concerning eligibility for health care benefits, the scope of health care benefits covered by the plan, and the rights and responsibilities of participants under the plan. The study engaged four expert readers to review the study sample and identify this information.
- The second stage measured empirically the readability level of the text, previously identified by the expert readers, that described the rules of the plan concerning eligibility, covered benefits, and participant rights and responsibilities under the plan. To measure readability levels, the study applied various standard readability assessment tests from the field of educational psychology to the selected text.

Description of the Study Sample—The study collected and tested the summary plan descriptions for 40 health care plans from a diverse geographic range of employers from across the United States, as well as large and small employers representing a variety of industrial sectors. The study sample included both health care plans sponsored by a single employer for its workers (single-employer plans), and health care plans sponsored jointly by several employers and a labor union for collective bargaining unit employees of the sponsoring employers (multi-employer plans). The study sample included different types of health care plans, including traditional indemnity health care plans offered through insurance companies, employer self-insured health care plans with a third party administrator and a utilization review manager, managed care plans, and one high-deductible health care plan with a corresponding health savings account feature.

The majority of the summary plan descriptions were obtained by conducting a random search of the Internet. Summary plan descriptions also were obtained from employees who participated in the plan. A few of the summary plan descriptions were obtained from attorneys who had represented plan participants in recent federal court litigation where the language of the summary plan description was related to the issue in dispute.

The Six Topic Areas Tested—The first stage of the study investigated whether expert readers reliably could identify important information in the summary plan descriptions concerning eligibility for health care benefits, the scope of health care benefits covered by the plan, and the rights and responsibilities of participants under the plan. The expert readers were instructed to identify information in the text that was relevant to the following six topic areas:

- Medical necessity clauses.
- *Firestone* clauses.
- Claim filing and appeal procedure clauses.
- Mental health and substance abuse benefits clauses.
- Pre-existing condition coverage exclusion clauses.
- Reimbursement or subrogation clauses.

Medical necessity clauses limit the benefits provided by the plan to medical treatment and health care services that are determined to be medically necessary by the plan's administrator. *Firestone clauses* affect a participant's legal rights under the plan by requiring the courts to defer to a plan administrator's decision to deny a claim for health care benefits under the plan, and to overturn that decision only if the court finds that the decision to deny the claim was arbitrary and capricious.³ *Firestone* clause language in a health care plan document is important because this language makes it more difficult for a participant to challenge a claim for plan benefits that has been denied by the plan administrator, such as a claim for benefits that has been denied based on a lack of medical necessity.

Claim filing and appeal procedure clauses describe a plan's internal administrative procedures. These procedures dictate the actions a participant must take to submit a claim for benefits and to appeal a claim for benefits that has been denied by the plan's administrator. *Mental health and substance abuse clauses* define the coverage of medical treatment and health care services for mental health or substance abuse conditions. These clauses typically impose additional procedural requirements that the participant must follow in order to obtain these benefits. *Pre-existing condition coverage exclusion clauses* describe the circumstances under which the plan excludes coverage for individuals who have a pre-existing health condition and define the duration of the exclusion period. *Reimbursement or subrogation clauses* allow the plan to recoup medical expenses paid by the plan that relate to a participant's injury that was caused by a third party. Reimbursement clauses require the injured participant to reimburse the plan if the participant later recovers against the third party who was responsible for causing the participant's injury.

Four expert readers participated in the first stage of the study. All four readers are considered experts on ERISA-regulated health care plans by their peers in the legal profession.⁴ The expert readers were asked to identify or "code" language in the summary plan description document according to the six topic areas selected for investigation. Two expert readers coded each summary plan description. The first author of the study coded all of the summary plan descriptions in the study sample. The three other expert readers each coded one-half, one-fourth, and one-fourth, respectively, of the study sample.

Each expert reader received detailed instructions that outlined the criteria to use in determining what language in the summary plan description should be identified as part of a topic area.⁵ For example, the instructions to the expert readers described medical necessity clause language as provisions that:

restrict the type of medical treatment that plans are obligated to cover. A medical necessity clause is an optional plan design feature that is used to reduce the cost of the health care plan by limiting the scope of coverage to treatment that is deemed to be medically necessary by the plan administrator.

The expert readers were instructed that for the topic area of mental health and substance abuse benefits, they should identify relevant language in the summary plan description document that described such items as:

- (1) the benefits and coverage limitations of the plan concerning treatment for mental health conditions or substance abuse; (2) the special procedural requirements of the plan, if any, for obtaining mental health or substance abuse benefits; and (3) special procedural requirements for the filing of claims or appealing denied claims for mental health or substance abuse benefits under the plan.

The expert readers were given further guidance concerning language that should be excluded from the study. For example, the expert readers were instructed to exclude the following information related to the topic area of claims filing and appeal procedure clauses:

(1) a description of the requirements of the plan for the pre-authorization of medical treatment, utilization review procedures, procedures for reviewing the appropriate length or continuation of a hospital stay, procedures for the coordination of benefits paid by multiple plans, or case management review procedures; (2) claim filing and appeal procedures that are contained in a *separate description* of prescription drug benefits, disability plan benefits, dental plan benefits, vision plan benefits, or other welfare plan benefits that are not medical benefits.

To measure reliability of identification, the study calculated inter-rater agreement among the expert readers. The first reader, who reviewed all of the summary plan descriptions, served as the standard for calculating agreement with each of the three other expert readers. The study calculated an agreement rate by taking the number of paragraphs that both expert readers marked as relevant to a particular topic area, and divided this number by the total number of paragraphs that either reader marked as relevant to a particular topic area.

To illustrate this methodology, assume that the first expert reader identified 10 paragraphs as part of the plan's medical necessity clause, and the second expert reader identified 9 paragraphs as part of the plan's medical necessity clause. The two expert readers agreed upon 8 paragraphs (i.e., they both marked off the same 8 paragraphs), but they disagreed on three others. For this summary plan description's medical necessity clause, the inter-rater agreement would equal 8/11, or 72 percent.

Application of Readability Assessment Tests—The second stage of the study assessed the readability of important language in the summary plan descriptions concerning eligibility for health care benefits, the scope of health care benefits covered by the plan, and the rights and responsibilities of participants under the plan. Rather than testing each summary plan description in its entirety for readability, the study assessed the readability of the text in the document that the four expert readers identified as relevant to the six topic areas. This methodology was used because not all of the information communicated through the summary plan description is equally important from the perspective of the plan participant. The six topic areas tested for readability focus on information that is fundamental to understanding how the health care plan works from the perspective of the plan participant.

In assessing the document language for readability, any text that had been previously identified by an expert reader as part of a topic area was tested for readability. This procedure guaranteed that the text from each summary plan description tested for readability was maximally inclusive and did not depend on the judgment of a single expert reader.

The selected text was submitted to the Flesch Reading Ease formula, the Flesch Grade Level formula, and the Fog Index. The Flesch Reading Ease formula yields a readability score between 0 and 100. Lower scores indicate the material is more difficult to comprehend (Flesch, 1948). The formula for the Flesch Reading Ease score takes into account average sentence length and average number of syllables per word. The resulting scores are associated with grade levels (e.g., 0–30 = college graduates; 30–50 = college years; 50–60 = 10th–12th graders).

The related Flesch Grade Level indicates the minimum education level required for the reader to be able to understand the document (Kincaid et al., 1975). The formula for the Flesch Grade Level also is based on average sentence length and average number of syllables per word. The Flesch Grade Level formula uses different coefficients from the Flesch Reading Ease Formula, and the output is stated in terms of grade level.

The Fog Index uses different indicators to measure language complexity. The Fog Index weighs the total number of words, words of three or more syllables, and sentences (Gunning, 1968). Commentary accompanying the Fog Index recommends that technical material should score no higher than 14, business material should score no higher than 12, and clerical material should score no higher than 8 (Thomas, Hartley, and Kincaid, 1975).

Results of the Study

Identification of Important Information—The first stage of the study investigated how reliably expert readers could identify important information contained in the summary plan descriptions. Figure 1 shows the inter-rater agreement rates, which are aggregated for all of the expert readers, for each of the six topic areas. The entries in this figure are in decimal format rather than in fractional or percentage format for ease of presentation.

Figure 1
Reliability of Identification as Measured by Inter-Rater Agreement

	Medical Necessity	Firestone	Claims & Appeals	Mental & Substance Abuse	Pre-existing Conditions	Reimbursement
Average Across Raters	0.42	0.64	0.83	0.69	0.75	0.87
Weighted Average	0.31	0.44	0.85	0.76	0.45	0.92

Source: Authors' calculations.

Row 1 of Figure 1 shows the average rate of agreement among the expert readers for each topic area. The unweighted averages shown in Row 1 ignore difference in the size of the text for a particular topic area. Because there was high variability in the size of the text assigned to a given topic area across plans, the study further calculated the weighted averages, which are shown in Row 2 of Figure 1. The weighted average calculation reflects an adjustment for variations in the size of the text (i.e., the number of paragraphs) for each topic area.

Figure 1 shows that regardless of the method of analysis, overall agreement varied widely across the six topic areas, ranging from 0.42–0.87 for an unweighted average and from 0.31–0.92 for a weighted average. The overall inter-rater agreement rate, averaged for all six topic areas, was 0.70 and 0.62, unweighted and weighted, respectively.

These results indicate that even expert readers experience some difficulty in reliably identifying important information in summary plan descriptions concerning eligibility for health care benefits, the scope of health care benefits covered by the plan, and the rights and responsibilities of participants under the plan. Some topic areas are substantially more difficult to identify reliably than are others. For example, inter-rater agreement was poor for medical necessity clauses and *Firestone* clauses under either method used. Inter-rater agreement also was poor for pre-existing condition clauses under the weighted average method. Inter-rater agreement was highest for reimbursement or subrogation clauses, and for claims filing and appeal procedure clauses.

This finding is potentially significant because the six topic areas investigated are fundamental to understanding how the health care plan works from the perspective of the plan participant. If even expert readers experience difficulty in identifying important information in the summary plan description, it is unlikely that the average plan participant, who is not an expert reader, will be able to do so.

Readability of Important Information—The second stage of the study performed a readability assessment for each of the six topic areas using the Flesch Reading Ease formula, the Flesch Grade Level formula, and the Fog Index. The readability assessment results were then subjected to various statistical tests. These statistical tests and their results are described first, followed by the readability assessment scores.

First, the study examined the alpha coefficients produced by performing the reliability analyses using standardized scores. Alpha coefficients are a statistical test that measures the lack of random noise among indicators. Higher alpha coefficients (>0.70) show that different types of measures agree with each other and are measuring the same construct. The alpha coefficients ranged from a low of 0.81 for pre-existing condition clauses to a high of 0.98 for claim filing and appeal procedure clauses. These results demonstrate that the three assessment tests of readability used in the study (the Flesch Reading

Ease formula, the Flesch Grade Level formula, and the Fog Index) were each measuring the same construct. This finding indicates that the data captured summary plan description readability.

Next, a single standard score was created that aggregated the readability test results for each of the six topic areas. This score was created separately for each of the three readability assessment tests used in the study. The study performed a statistical test, known as analysis of variance, to determine if there were differences in readability among the six topic areas investigated for each of the three measures of readability used. An analysis of variance tests whether there are statistically significant differences (beyond chance) among a group of sample means on a specific variable or measure. The analysis of variance showed that readability levels did not differ across the six topic areas for the three readability assessment tests used in the study.⁶

Row 1 of Figure 2 below presents the aggregated average readability scores for the six topic areas according to the type of readability assessment test used. Figure 2 shows entries for all six topic areas combined because the analysis of variance demonstrated no differences in readability across the six topic areas.

Figure 2
Readability Assessment Measures

Flesch Reading Ease	Flesch Grade Level	Fog
Mean = 37.66	Mean = 13.23	Mean = 16.37
Standard Deviation = 6.89	Standard Deviation = 1.54	Standard Deviation = 1.63
Range: 26–55	Range: 9.35–15.94	Range: 10.90–19.06
95% confidence: 35.46–39.86	95% confidence: 12.74–13.72	95% confidence: 15.85–16.89
Source: Authors' calculations.		

The Flesch Reading Ease mean of 37.66 indicates that the language tested is written at a college reading level. The related Flesch Grade Level indicates that the minimum education level required for the reader to be able to read the language used is 1+ year of college (i.e., 13th grade). The Fog Index mean of 16.37 is higher than the recommended readability level for technical material (14), or business material (12) (Thomas, Hartley, and Kindcaid, 1975).

To be able to generalize the results from the 40 summary plan descriptions in the study sample to the population of summary plan descriptions as a whole, it was necessary to perform a series of additional statistical tests. First, the study calculated two measures of variability in the readability of the tested language, the *standard deviation scores* and the *range scores*. A standard deviation score shows the average amount of variability for the particular readability test used. A range score identifies the lowest and highest readability levels in the study sample for the particular type of readability test used. Rows 2 and 3 of Figure 2 show the results from further statistical testing to determine the standard deviation scores and the range scores.

The standard deviation scores indicate that there was a wide range of readability for the six topic areas across the study sample. The relatively high standard deviation scores are consistent with the range scores. Using the Flesch Grade Level indicator, for example, the range scores indicate that some of the summary plan descriptions in the study sample used language that was written at a 9th grade reading level. Other summary plan descriptions in the study sample used language that was written at nearly a college graduate (16th grade) reading level. The results of the standard deviation scores and the range scores across all three tests of readability used were consistent, which indicates that the standard deviation scores are reliable.

Finally, to be able to generalize the results from the study sample to the population of summary plan descriptions as a whole, the study calculated 95 percent confidence intervals. The 95 percent confidence interval is a statistical procedure used to estimate where the true mean lies for the population from which a sample is drawn. Confidence intervals depend upon sample size and standard deviation scores. Therefore, larger or smaller samples, and larger or smaller standard deviation scores, could produce slightly different 95 percent confidence intervals. Row 4 of Figure 2 shows the results of the 95 percent confidence level calculations.

For this particular study sample, given the observed sample mean of 37.66 on the Flesch Reading Ease scale, the 95 percent confidence level shows that we can be 95 percent confident the mean for the

population represented by the study sample lies between 35.46 and 39.86. In other words, the 95 percent confidence interval results indicate we can be 95 percent confident that the true population mean of the Flesch Reading Ease scale falls within the upper and lower boundaries that were calculated (here 35.46 and 39.86) using the study sample.

Finally, the study examined whether readability levels differed between summary plan descriptions for health care plans sponsored by single employers and multi-employer health care plans, which are sponsored for workers who are members of a collective bargaining unit that is represented by a labor union. The study found that readability levels using all three tests for readability did not show statistically significant differences between single-employer plans and multi-employer plans.

Future Research Implications of the Study Results

In considering the implications of the study results, it is important to note two limitations of the readability assessment tests used in the study. First, because the formulas used by the tests consider both sentence length and number of multi-syllable words, it is impossible to determine the extent to which the use of medical terminology or jargon in the summary plan description documents may be affecting the results of the readability assessment tests. Second, readability assessment tests are not designed to predict objective comprehension by individual plan participants (Black, 1981; Duffy and Kabance, 1982). In other words, an individual may be able to read the written text and yet not *comprehend* the information the text is intended to communicate. Additional research is needed to measure directly how well individual participants comprehend important information in a summary plan description. Such additional research would assess an individual's level of knowledge about the plan after reading the summary plan description.

Nevertheless, the study results paint a troubling picture. The results of the study suggest that, based on the difficulties encountered by the expert readers, important information conveyed through a summary plan description may be difficult for the average plan participant to identify reliably. Once identified, the language used to convey important information in the summary plan description document may be written at a level that is too high for its intended audience—the “*average* plan participant”—to understand. In other words, the primary communication tool used to provide important information to workers who participate in their employer's health care plan often may be unreadable to them.

The findings of the study corroborate and support the testimony collected by the Working Group on Health and Welfare Benefit Plans' Communications that many workers today may not be receiving information about their health care plans that is understandable to them. The study results confirm the Working Group's conclusion that employers and plan administrators face significant challenges in communicating important information about their health care plans through written summary plan descriptions (U.S. Department of Labor, 2005).

The study findings also have potential implications for the projected future effectiveness of consumer-directed health care plans. To make informed utilization decisions, participants in consumer-directed health care plans need to understand how the plan operates, the benefits covered by the plan, and their rights and responsibilities under the plan. The results of the study provide quantifiable evidence that many participants are unable to understand the primary method used by employers and plan administrators to communicate the essential information in consumer-directed health care plans using only written summary plan descriptions. At the very least, these results suggest that fundamental improvements are needed in the readability of written summary plan descriptions, and that employers and plan administrators should explore the use of alternative methods of communication to plan participants beyond the written SPD.

Bibliography

Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs, American Medical Association. "Health Literacy: Report of the Council on Scientific Affairs." *Journal of the American Medical Association*. Vol. 281, no. 6 (February 1999): 552–557.

Black, Bernard. "A Model of Plain Language." *Stanford Law Review*. Vol. 33 (January 1981): 255–300.

Duffy, Thomas M., and Paula Kabance. "Testing a Readable Writing Approach to Text Revision." *Journal of Educational Psychology*. Vol. 74, no. 5 (1982): 733–748.

Flesch, Rudolf. "A New Reliability Yardstick." *Journal of Applied Psychology*. Vol. 32, no. 3 (June 1948): 221–233.

Fronstin, Paul, ed. *Consumer-Driven Health Benefits: A Continuing Evolution?* Washington, DC: Employee Benefit Research Institute, 2002.

_____. "Health Savings Accounts and Other Account-Based Health Plans." *EBRI Issue Brief* no. 273 (Employee Benefit Research Institute, September 2004).

Fronstin, Paul, and Sara Collins. "Early Experience with High-Deductible and Consumer-Driven Health Plans: Findings From the EBRI/Commonwealth Fund Consumerism in Health Care Study." *EBRI Issue Brief* no. 288 (Employee Benefit Research Institute, December 2005).

Gazmararian, Julie A., et al. "Health Literacy among Medicare Enrollees in a Managed Care Organization." *Journal of the American Medical Association*. Vol. 281, no. 6 (February 1999): 545–551.

_____. "Literacy and Knowledge of Chronic Disease." *Patient Education & Counseling*. Vol. 51 (2003): 267–275.

Gunning, Robert. *The Technique of Clear Writing*. New York: McGraw-Hill, 1968.

Hibbard, Judith H., and Jacquelyn J. Jewett. "Will Quality Report Cards Help Consumers?" *Health Affairs*. Vol. 16, no. 3 (May/June 1997): 218–228.

Hibbard, Judith H., et al. "Can Medicare Beneficiaries Make Informed Choices?" *Health Affairs*. Vol. 17, no. 6 (November/December 1998): 181–193.

Isaacs, Stephen L. "Consumers' Information Needs: Results of a National Survey." *Health Affairs*. Vol. 15, no. 4 (Winter 1996): 31–41.

Kincaid, J. Peter, R. Fishburne, R.L. Rogers, and B.S. Chissom. *Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel*. Springfield, VA: National Technical Information Service, 1975.

Thomas, Georgelle, R. Derald Hartley, and J. Peter Kincaid. "Test-Retest and Inter-Analyst Reliability of the Automated Readability Index, Flesch Reading Ease Score, and the Fog Count." *Journal of Reading Behavior*. Vol. 7, no. 2 (Summer 1975): 149–154.

Vastag, Brian. "Low Health Literacy Called a Major Problem." *Journal of the American Medical Association (JAMA)*. Vol. 291, no. 18 (May 12, 2004): 2181–2182.

U.S. Department of Education. National Center for Education Statistics. Institute for Education Sciences, 2003. *National Assessment of Adult Literacy* (December 15, 2005), <http://nces.ed.gov/naal/>).

U.S. Department of Labor. Employee Benefits Security Administration. *Report of the Working Group on Health and Welfare Benefit Plans' Communications*. Washington, DC: U.S. Department of Labor, November 2005.

Williams, Mark V., et al. "Inadequate Functional Health Literacy at Two Public Hospitals." *Journal of the American Medical Association*. Vol. 274, no. 21 (December 1995): 1677–1682.

Endnotes

¹ Colleen E. Medill is an EBRI Fellow and professor of law at the University of Nebraska-Lincoln College of Law. Richard L. Wiener is professor of psychology, director of the University of Nebraska Law and Psychology Program, and courtesy professor of law at the University of Nebraska-Lincoln. Brian H. Bornstein is professor of psychology and courtesy professor of law at the University of Nebraska-Lincoln. E. Kiernan McGorty is a graduate of the University of Nebraska-Lincoln College of Law and a graduate student in the doctoral program in psychology at the University of Nebraska-Lincoln. This research study was jointly funded by the Employee Benefit Research Institute and The Commonwealth Fund.

² U.S. Department of Education, National Center for Education Statistics, 2003 National Assessment of Adult Literacy (Institute for Education Sciences, Dec. 15, 2005, <http://nces.ed.gov/naal/>).

³ Firestone clauses originate from a 1989 United States Supreme Court decision, *Firestone Tire & Rubber Co. v. Bruch*, 489 U.S. 101 (1989). In *Firestone*, the Supreme Court established the judicial principle that the federal courts must give deference to a plan administrator's decision to deny a claim for health care benefits under the plan. This judicial principle applies only if the health care plan expressly gives the plan administrator the discretionary authority to determine eligibility for benefits or to construe the terms of the plan.

⁴ All of the expert readers are lawyers who specialize in ERISA. One of the expert readers (Medill, also the first author of the study), is a lawyer and law school professor with a combined total of 14 years of private legal practice and academic research experience with ERISA. Two of the other expert readers each had six years of ERISA legal practice experience. The fourth expert reader had three-and-a-half years of ERISA legal practice experience and 10 years of experience as a human resources manager responsible for the administration of a health care plan sponsored by an employer.

⁵ The instructions that the researchers provided to the expert readers are available from the first author of the study. Correspondence concerning the study or requests for a copy of the instructions may be sent to cmedill2@unl.edu

⁶ The results of the analysis of variance test were $F(5,55) = 1.26, p = 0.29$.