Alan G. Sawyer, Juliano Laran, & Jun Xu

The Readability of Marketing Journals: Are Award-Winning Articles Better Written?

This is a study of the readability of articles in four marketing journals: *Journal of Marketing*, *Journal of Marketing Research*, *Journal of International Marketing*, and *Journal of Public Policy & Marketing*. For each journal, the authors compare articles that have won an award with articles that have not. The authors find that award-winning articles are more readable, as measured by indexes focusing on sentence and word length, than nonwinning articles. The authors also identify and analyze other characteristics of more readable journal articles and discuss the importance of good writing.

Keywords: readability, award-winning articles, marketing journals, path analysis

The October 2005 *Journal of Marketing* (*JM*) includes 11 essays about challenges to marketing and the content and impact of published research in marketing journals. The essay "Readability and the Impact of Marketing" (Bauerly, Johnson, and Singh 2005) documents a steep decline of readability in *JM* articles in the early 1970s, when marketing academics became more scientific and specialized. Since then, readability has remained low. This article offers an expanded empirical investigation of the readability of marketing journal articles.

The current manuscript evaluation criteria for *JM* make it clear that readability is important. "All manuscripts are judged not only on the depth and scope of the ideas presented and their contributions to the field, but also on whether they can be read and understood. Keep sentences short so the reader does not get lost before the end of a sentence.… The journal is designed to be read, not deciphered" (American Marketing Association [AMA] 2005).

Despite frequent calls for improved readability, many marketing journal editors view writing in academic journals as badly in need of improvement. As Winer (1998, p. iv) states in a *Journal of Marketing Research* (*JMR*) editorial, "Authors consistently underestimate the importance of writing to a paper's ultimate disposition. Reviewers are justifiably impatient with manuscripts in which the contribution is not clear, there are many typos (even in this era of spelling checkers), and so on. Do not believe that these kinds of problems will be eliminated through revisions; if the writing is sufficiently unclear, the author might not get that opportunity." Finally, in his outgoing *Journal of Consumer Research* editorial, Lutz (1990, p. 244) states, "Frankly, it is appalling to see how badly so many of us write." In addition to bad writing, virtually all editors focus on the excessive length of submitted manuscripts. "In the vast majority of cases, manuscripts are simply too long relative to their potential contribution.… Ceteris paribus, the shorter the article, the more likely it is that someone will actually read it" (Lutz 1990, p. 245).

Clear scientific writing is essential for research to be understood, published, and impactful. As Mick (2005, p. 1) states, "The unavoidable reality is that writing effectively is not just a necessary condition for getting published. Better writing propels an academic's influence and reputation, and the finest writing is more often found among the most distinguished researchers." Moreover, good writing is essential "if marketers want to communicate across specializations, across functional areas, and outside the marketing discipline and if marketing science is to influence practice" (Bauerly, Johnson, and Singh 2005, p. 19).

Research Objectives

The prime research objective was to test the validity of Mick's (2005) logical speculation that impactful articles written by the most distinguished researchers are more readable. Readability is defined as "the ease of understanding or comprehension due to the style of writing" (Klare 1963, p. 15) and "the extent to which (a group of) readers understand it, read it at an optimal speed, and find it interesting" (Dale and Chall 1949, p. 19). We used articles that

Alan G. Sawyer is J.C. Penney Professor of Marketing (e-mail: alan.sawyer@cba.ufl.edu), and Juliano Laran (e-mail: laran@ufl.edu) and Jun Xu (e-mail: jun.xu@cba.ufl.edu) are doctoral candidates in Marketing, University of Florida. The senior author thanks the Warrington College of Business and the J.C. Penney Foundation for financial assistance with this research. Roger Calantone of Michigan State University and Dwayne Ball and Daniel Chen of the University of Kansas provided valuable assistance with the principal components analysis and path analysis, for which the authors are indebted. Taking their own advice, the authors asked Lita Kurth to copyedit one of their "final" drafts, and they are grateful for her improvements to the manuscript. They also appreciate the helpful advice of the anonymous *JM* reviewers. All three authors contributed equally to this work.

ISSN: 0022-2429 (print), 1547-7185 (electronic)

Journal of Marketing
Vol. 72 (January 2008), 108–117

had won a major award in an AMA academic journal as the benchmark of distinguished research. Another research objective was to identify article and author characteristics that may contribute to or at least correlate with readability. Bauerly, Johnson, and Singh (2005) speculate that a reason for the decline in readability is that more articles are written by academics from more methodologically rigorous and theoretically based doctoral programs. This speculation seems valid, and several other aspects of the authors' backgrounds and the content of the articles may be related to better readability. Because Bauerly, Johnson, and Singh did not code the contents of analyzed articles, they were unable to examine correlates of readability other than the publication date.

In our research, we code the content of sampled articles and test the relationship of readability to various article types and author characteristics. We use these characteristics to control statistically for differences among award-winning and nonwinning articles. The content of the article and other such characteristics might influence readability. We predict better readability for some types of empirical articles than nonempirical ones for three reasons. First, empirical articles, particularly experiments, have a prototypical report format that might enable briefer, more readable sentences. Second, nonempirical articles focusing on managerial, theoretical, or other complex issues might contain longer and more complex sentences with longer words. A particular type of nonempirical article, the literature review, might be particularly less readable. Third, empirical articles with data and statistics might be rated as more readable because most contain more numbers. This could result in an artifactual relationship because numbers are coded as one-syllable words with only one or a few characters. We had no expectation regarding whether different journals would be rated as more readable, other than the variation in content or author experience. Articles in more prestigious academic journals might be more readable if, for example, authors write more carefully and seek more advice for manuscripts targeted at these journals. However, articles in journals with a more practitioner-oriented audience may be more readable if, in consideration of those readers, authors write with simpler sentences and less jargon.

Author characteristics may be important. For example, more experienced authors likely write better, regardless of whether the article has won an award. Top researchers likely think more clearly, which is an essential skill for good writing. Award-winning authors are likely to be more experienced and have an advantage of more practice and more feedback from editors and colleagues on more manuscripts. Coauthorship is also possibly related to readability. Coauthored articles might benefit in terms of readability because each coauthor would try to improve the writing of the other authors or would work on assigned sections about which they are more knowledgeable and thus write better. Conversely, coauthoring may lead to coordination problems that could result in less clear, more cumbersome writing. Hartley, Pennebaker, and Fox's (2003) literature review shows mixed results for the readability of coauthored articles, and their own empirical investigation finds no differences between single-authored and coauthored works.

Characteristics of Good Writing

Scientific writing presents several challenges. Complex topics may require complex writing to convey accurate information, with frequent use of jargon, qualifying phrases, data, statistics, footnotes, and references. In addition to demanding a correct presentation of all the facts, readers "won't put up with being delayed, lost, confused, bored or taken down unnecessary trails by failures of craft" (Zinsser 1988, pp. 57–58). The many dimensions of good scientific writing include having good grammar and sentence structure; using appropriate, short, and easy-to-understand words and sentences; avoiding unnecessary technical jargon; and keeping the length proportionate to the likely contribution. Other imperatives beyond the basics are probably more important and subtle (Sawyer 1988). These include properly positioning and "selling" the contribution of the research (Peter and Olson 1983), writing for a specific audience and journal, and using an interesting writing style with a clear and appropriate structure (Holbrook 1986).

Although existing objective methods do not measure the more subtle aspects of good writing, several formulas that calculate readability indexes are available to assess the basics (for a summary of this research, see DuBay 2004). These formulas include the length of words, as expressed by the number of syllables or characters, and the length of sentences, as expressed by the number of words or long words. These basics do not guarantee a readable manuscript; the formulas "measure only 'surface features' of text and ignore content and organization" (DuBay 2004, p. 35). However, although sentence and word length do not *cause* readability difficulty, they are highly *correlated* with this problem and thus can be used to suggest possible improvements.

Attention to readability indexes can help authors improve their writing. As Judith Swan (qtd. in Knight 2003, p. 377) stated, "Jargon is less pernicious if you understand what is going on." The structure of the sentence is more apt than long words to be the prime readability problem (Gopen 2004; Gopen and Swan 1990). Readers persist longer with easier-to-read material and find it more interesting if the writing matches their preferred reading level.

Readability indexes are not without problems (Redish and Selzer 1985). For example, shorter sentences and words are not always clearer than longer ones, and readability does not measure the appropriateness or accuracy of the content. Moreover, although scientific texts with many scientific terms may have lower readability ratings, they may be easier to read for those trained in subspecialties, who may even prefer jargon or are bored with simpler, "more readable" text.

Despite these issues, readability formulas have received corroboration from initially skeptical cognitive psychology researchers, who have tried to go beyond simple measures of word and sentence arithmetic. Cognitive theorists who study how readers learn posit that readers construct meaning because meaning is not in the words by themselves on the page. Readers store information in long-term memory, link new information to prior knowledge, and use metacogniton to think about and plan the learning process (DuBay 2004; Gopen and Swan 1990). After examining the empiri-

cal effects on comprehension of cognitive and structural text elements, such as coherence and organization, several cognitive theorists have reversed their negative opinions about readability formulas (see DuBay 2004, p. 35). For example, Kintsch and Miller (1984, p. 222) conclude that "these formulas are correlated with the conceptual properties of text" and are the strongest predictors of reading speed and comprehension difficulties.

Readability of Top Journals and Articles

Bauerly, Johnson, and Singh (2006) compared the abstracts of 70 *JM* articles with 200–400-word sections of their introductions. The abstracts were rated significantly more readable than the material from the articles, though the two scores were significantly correlated. Bauerly, Johnson, and Singh concluded that the authors knew how to write better than they did in the text. They speculated that a reason for the poorer writing in the text is that some researchers might write less simply than desirable and overuse technical jargon to impress readers. Armstrong (1980) conducted two tests of this possible phenomenon. In the first test, he found that the readability was inversely related to ratings of journal prestige. In a second, better-controlled test, he took four excerpts from marketing journals and manipulated their reading difficulty by changing the length of sentences, varying the difficulty of words, and eliminating/adding unnecessary words. Thirty-two faculty members judged the less readable excerpts with the same content to be more competent. Hartley, Trueman, and Meadows (1988) found the same result with psychology articles. Armstrong (1989) subsequently reanalyzed both data sets and concluded that more easily read articles were rated less prestigious ("If I can read this easily, it must not be prestigious"), but beyond some point, differences in readability had no further effect.

Additional research has found counterintuitive relationships between readability and factors such as the number of citations of the article and rejection of a manuscript by a journal. Stremersch, Verniers, and Verhoef (2007) analyzed four top-rated marketing journals and found a small but statistically significant, negative relationship between readability and the number of times an article had been cited. They concluded (p. 1) that "the readability of an article may hurt academic credibility and impact." Articles with fewer citations were more readable. A similar negative relationship to readability was reported by Metoyer-Duran (1993), who analyzed the readability scores of articles accepted and rejected for publication in a library science journal. Readability scores of rejected manuscripts were significantly better than those of accepted ones. In contrast to the negative results of Armstrong (1980) and others, Oppenheimer's (2005) experimental research showed greater preference for more readable text. He conceptualized that because of greater processing fluency, texts with simpler words should result in more positive judgments. The results of five highly controlled experiments were consistent with this prediction and his hypothesized explanation. Respondents judged the authors of texts with simpler words to be more intelligent, and this effect was mediated by ratings of processing fluency. Direct manipulations to increase the fluency of texts resulted in judgments that were more positive. Only when the source of fluency was obvious did respondents discount their reliance on the fluency cue.

Additional correlational evidence from research in psychology suggests that articles that are considered more important are more readable than their less important counterparts on the standard readability indexes. Hartley, Sotto, and Pennebaker (2002) reported four correlational studies that took an excerpt from the introductory pages of each article and scanned it into a word-processing package. The first study compared the readability scores for excerpts from 19 of the most influential journal articles in cognitive science with 19 articles that were not cited as influential but were published in the same journal issues. The impactful articles had significantly better readability scores, and there was a positive .41 rank correlation between the Flesch scores and the position of 35 articles within the 100 most influential publications (which also included books and monographs). Hartley, Sotto, and Pennebaker's second study found that 36 articles from *Forty Articles That Changed Psychology* (Hock 1992) had significantly better readability scores and shorter sentences than 36 counterpart articles not cited by Hock. However, two other studies by Hartley and colleagues found no relationship between readability and the number of citations.

In summary, the evidence is mixed regarding whether articles written in a less readable way are preferred, more likely to be accepted at a journal, associated with greater impact, considered more prestigious, or written by more intelligent authors. Most relevant for the purposes of this study is the evidence that renowned articles may be more readable.

Research Method

This research assessed whether award-winning marketing articles in four academic journals published by the AMA are more readable than counterpart articles that have won no awards. We also examined variables that may characterize more readable publications. To avoid problems with the validity of sampling only small, single excerpts from articles and avoid errors between scanned and original texts, we copied almost all of each article's text downloaded from Internet journal databases available in HTML format into Microsoft Word. Unlike prior research, our study used multiple measures of readability.

Sample

Journals. The sample included award-winning articles in *JM*, *JMR*, *Journal of International Marketing* (*JIM*), and *Journal of Public Policy & Marketing* (*JPP&M*). There were four reasons for choosing this particular set of journals. First, these academic journals are the only ones published by the AMA. Second, online versions of their articles published since 1990 are available in HTML format. Third, these journals differ considerably and thus provide within-study replication across a diverse set of journals. In particular, *JPP&M* and *JIM* vary greatly from *JM* and *JMR* in terms of their focus, impact, and prestige. The *Web of Sci-*

ence Journal Citation Reports impact index (based on the number of citations of an average article in a journal within two years of publication) and Baumgartner and Pieters's (2003) structural influence index (based on the proportion of citations in other journals of all articles published in that journal) show that the recent impact and overall influence were high for both *JM* (4.1 and 19.1, respectively) and *JMR* (2.6 and 16.4) and lower for both *JPP&M* (1.2 and .8) and *JIM* (.4 and .2). Moreover, academics vary widely in their rankings of these journals. Two studies (Bauerly and Johnson 2005; Hult, Neese, and Bashaw 1997) show large differences in the ranking of *JM* (1st in both studies) and *JMR* (3rd and 2nd, respectively) compared with *JPP&M* (36th of 54 journals and 12th of 34 marketing journals) and *JIM* (unranked and 26th of 34).

The fourth, and most important, reason for choosing these four journals is that each presents annual awards for outstanding articles. *Journal of Marketing* gives two awards for articles published during the previous year: the Harold H. Maynard Award for the best article on "marketing theory and thought" and the Marketing Science Institute/H. Paul Root Award for the "best contribution to marketing practice."[1] *Journal of Marketing Research* also gives two awards: the William F. O'Dell Award for the article published five years previously that made "the most significant, long-term contribution to marketing theory, methodology, and/or practice" and the Paul E. Green Award for the article published during the previous year that "shows or demonstrates the most potential to contribute significantly to the practice of marketing research and research in marketing." *Journal of International Marketing* also gives two awards: the S. Tamer Cavusgil Award for an article from the previous year that made "the most significant contribution to the advancement of the *practice of international marketing management*" and the Hans B. Thorrelli Award for the article published within the past five years that has made the "most significant, long-term contribution to *international marketing theory or practice*." *Journal of Public Policy & Marketing* gives the Thomas C. Kinnear Award for "significant contribution to the understanding of marketing and public policy issues" within the past three years.

Nonwinning articles. For each award-winning article, we chose a counterpart nonwinning article from the same journal issue to control for temporal issues, such as editors and reviewers. We arbitrarily set the nonwinner as the penultimate article in the issue; we considered the last article the one immediately preceding any notes or rejoinders. The reason for choosing nonwinning articles positioned near the end of the issue was to attempt a stronger "manipulation" of article quality. Many editors place articles they view as more important toward the front of the issue (e.g., Medoff 2003). If the beginning article was an editorial, we coded the article immediately following any editorials as the first. In some years, there were multiple recipients of a given award or articles that received multiple awards. We counted articles that won multiple awards (e.g., both the Green and the O'Dell awards in *JMR* in a given year) only once with only one counterpart nonwinner. If different award-winning articles appeared in the same issue, the article appearing just before the penultimate one became the second nonwinner. For example, if there were 11 articles in an issue containing an award winner and the final 3 were research notes, the nonwinner would have been the seventh article. If the same issue contained a second award-winning article, its nonwinning counterpart would have been the sixth. The final sample of 81 award-winning articles included 32 from *JM*, 22 from *JMR*, 15 from *JIM*, and 12 from *JPP&M*. These articles included 13 Maynard Award winners, 15 MSI/Root Award winners, and 4 that had won both awards in *JM*; 11 O'Dell Award winners, 9 Green Award winners, and 2 that had won both awards in *JMR*; 7 Thorrelli Award winners and 8 Cavusgil Award winners in *JIM*; and 12 Kinnear Award winners in *JPP&M*. If we include the 81 nonwinners from the same issues as the winners, the sample totaled 162 articles.

Measures

After we copied each entire article into Microsoft Word, we trimmed the text of each article of the title and abstract; all headings and subheadings; any mathematical equations (note that some HTML versions we used did not include the full text of any equations); quotations that included more than one sentence; and all tables, figures, and footnotes. We uploaded the remaining text to a Web site (Taylor and Intuitive Systems 2006) that calculated several indexes of readability that include word length defined in various ways (e.g., characters or syllables per word, words longer than a certain number of syllables or characters) and the length of sentences (number of words or words of a particular length per sentence). Four readability scores, normed to the school grade appropriate for that reading difficulty, were the Flesch–Kincaid (hereinafter, the Kincaid score), the automated readability index (ARI), the Gunning–Fog index (Fog), and the SMOG grade-level scores. We also calculated the Flesch index and the Lix index, which vary between 1 and 100. (The Appendix contains the formulas for all indexes.) Other readability measures produced by the Taylor and Intuitive Systems (2006) Web site included the number of characters in the manuscript, the number of words, their average length in terms of the number of characters and syllables, the number of sentences, and their average length.

In addition to the individual indexes, we calculated an average score, which we labeled as "average readability grade," from the Kincaid, ARI, Fog, and SMOG scores (coefficient $\alpha = .98$). Our readability results focus on the average readability grade level and three individual readability scores: the Flesch index (used by Bauerly, Johnson, and Singh 2005; Hartley, Pennebaker, and Fox 2003), the Lix score, and, though it is included in the average readability grade score, the Kincaid grade score presented by itself, because some readability researchers consider it the most appropriate one for technical documents (DuBay 2004). We

[1]A third *JM* award was established in 2001 by the Sheth Foundation for the *JM* article published 6–10 years prior with the largest contribution to marketing. Our analysis did not include Sheth Foundation Award articles, because all but one also won one of the two other *JM* awards.

also report other, perhaps more interpretable measures, including the length and number of sentences and words.

Other variables in the data set were potential correlates of readability. These included the number of authors, the award, the publication year, the issue number, the position of the article in the issue, the number of references within the article, the number of citations of the article, and its status as either an award winner or nonwinner. The number of references within the article may serve as a surrogate for a literature review article. To control for the length of the manuscript, which might be correlated with the number of references used, we divided the number of references within the article by the number of sentences. We obtained information about publications and citations from the Web of Science Social Science database. We coded the number of publications in six marketing journals (the four AMA journals included in our research plus *Journal of Consumer Research* and *Marketing Science*) for the most published author and the combined total publications of all authors. We divided the number of citations of each article by the number of years in print to create an annual citations variable.

Finally, we coded three other article characteristics. The first was whether the article included equations before being trimmed for analysis. Of the articles, 36% contained equations in the original text. A second variable, data source, included subcategories of data from secondary sources (30%), from experiments (14%), or from authors' surveys (41%). In addition, 26% were nonempirical and thus contained no data.[2] A third variable category included codes for whether the article topic was methodological (25%), managerial (57%), behavioral (17%), or theoretical (14%). We coded these subcategories as dummy variables. Some categories within both the data source and topic variables were highly negatively correlated. Even after we dropped one category so that any analysis would not be overidentified, some subcodes, such as methodological and managerial, remained nearly mutually exclusive. Thus, we kept only the data source and topic classifications with small, nonsignificant intercorrelations. The final set of five article characteristics included whether the article (1) contained equations, analyzed data from (2) experiments or (3) secondary sources, and focused on (4) methodological or (5) behavioral issues.

We then factor-analyzed these five article characteristics with the recommended M-plus algorithm (Muthén and Muthén 2004), which calculates tetrachoric correlations for a confirmatory factor analysis (CFA) for dichotomous variables instead of Pearson phi correlations. The results were two orthogonal principal components. The first component, which we labeled as "behavioral experiments," had a .91 correlation with the behavioral and experimental variables, and the second component, which we labeled as "methodological secondary data models," had a .83 correlation with each of the other three article characteristics. The fit of the CFA model was excellent (χ^2 = 8.32, d.f. = 5; comparative fit index = .98; and root mean square error of approximation = .064). We used these two components in further analyses that included descriptions of article characteristics.

[2]Because some articles were coded into more than one category, these percentages add to more than 100%. For example, one article, which analyzed data from field experiments previously conducted over many years by a commercial research firm, was coded in both the experiment and secondary data categories.

Results

Overall Readability

The average readability grade score across all articles averaged 16.2, which is considered difficult reading for people below the grade of a college senior. In comparison, the average readability grade levels for the *New York Times* and *PC World* are 7.6 and 11.7, respectively (Taylor and Intuitive Systems 2006). The Flesch readability score, which has been widely used in prior research and is the only index for which higher scores indicate better readability, averaged 35.3 for these 162 journal articles, compared with average scores of 70.4 and 59.4, respectively, for the *New York Times* and *PC World* (Taylor and Intuitive Systems 2006). Of the 15 articles with the best readability grade scores, 13 were winners, with scores ranging from 12.3, suitable for a high school senior, to 14.4, suitable for a sophomore in college. Of the 11 articles rated least readable, 9 were nonwinners, with the readability grade scores ranging from presumably graduate school levels of 21.3 to 18.3. Table 1 shows the scores for the 5 articles rated most and least readable.

Main Results

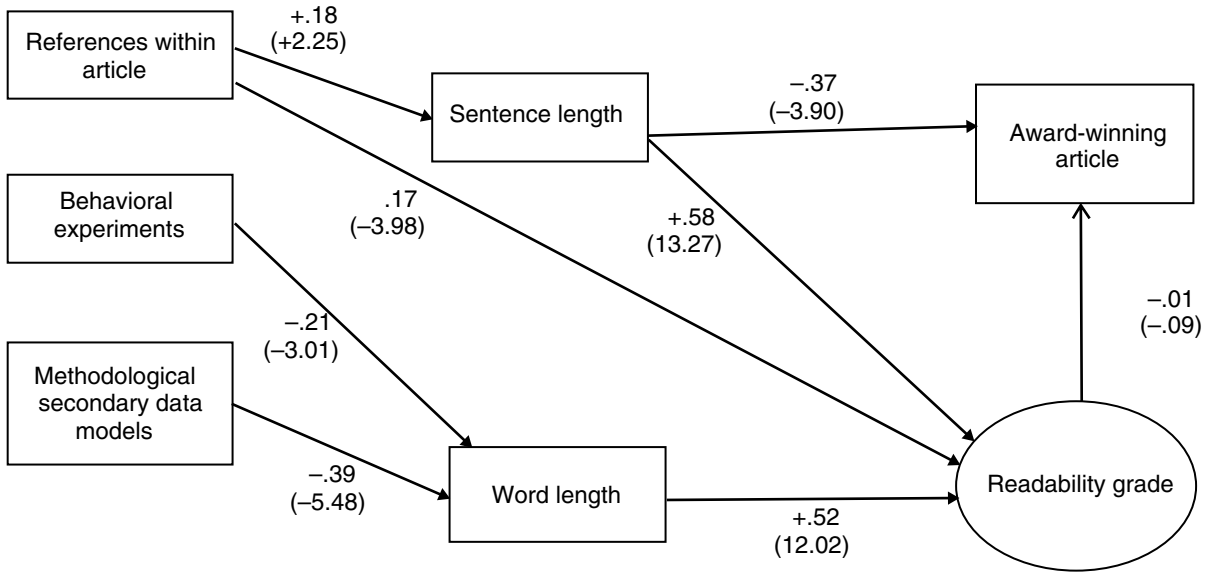
The prime goal of this research was to test whether award-winning articles in marketing journals were more readable than nonwinners. We used readability grade, as well as other variables, to define readability. We also analyzed the relationships between winning articles and the length of words and sentences. We defined word length as the number of characters and sentence length as the number of words. The second research objective was to search for any article characteristics in addition to word and sentence length that distinguish more readable articles. Specifically, we examined three article characteristics: the two principal component scores from the CFA (behavioral experiments and methodological secondary data models) and the number of references per sentence for each article.

To test these two objectives and assess various relationships simultaneously, we used a path analysis that examined the relationships among single measures of each of seven variables: articles with behavioral experiments and articles focusing on methodological secondary data models, the number of references per sentence, the length of both words and sentences, readability grade, and whether the article had won an award. The results of the path analysis, which we performed with LISREL, showed an excellent fit (χ^2 = 12.35, d.f. = 10, $p < .25$; comparative fit index = .99; and root mean square error of approximation = .039). The respective standardized beta coefficients and z-values appear in Figure 1.

TABLE 1
Articles Rated Most and Least Readable by Average Readability Grades

Authors	Award Winner	Readability Grade	Flesch Score	Sentence Words
Most Readable				
Golder and Tellis (1993)	Yes	12.3	50.7	18.2
Desiraju and Shugan (1999)	Yes	12.5	53.0	20.3
Stremersch and Tellis (2002)	Yes	13.2	47.5	19.8
Robinson and Min (2002)	No	13.6	44.0	19.5
Hoch, Dreze, and Purk (1994)	Yes	13.7	48.0	22.3
Least Readable				
Vorhies and Morgan (2003)	No	21.3	11.4	31.1
Prasad, Ramamurthy, and Naidu (2001)	No	20.2	18.5	31.5
Morgan, Kaleka, and Katsikeas (2004)	No	19.7	19.4	30.1
Glazer (1991)	Yes	19.6	24.9	28.4
Gencturk and Kotabe (2001)	No	19.3	20.8	29.5

Figure 1
Causal Path of Article Characteristics on Readability Grade



Notes: All coefficients are statistically significant at $p < .05$, except for the null relationship between "Readability" and "Award-winning article." Because lower grades indicate better readability, a negative coefficient indicates a positive relationship to readability. For example, longer words and sentences cause higher readability grades (i.e., poorer readability). The top value beside each path is the size of the standardized relationship. The resulting Z-test is in parentheses.

Award winners versus nonwinners. An analysis of covariance (ANCOVA) with the three leftmost covariates in Figure 1 and journal and winner as fixed factors augmented the path analysis. The ANCOVAs indicated that every readability index in the Appendix was significantly better for award-wining articles ($p < .02$). Word length was related to readability but not to award-winning status. Sentence length was related to both readability and whether the article won an award.

However, when we controlled for all of the variables in the model simultaneously in the path model, there was no direct relationship between readability grade and whether an article won an award. As Figure 1 shows, the path analysis indicated that sentence length was the only component of readability grade that distinguished award winners from nonwinners. This was presumably due to the fact that the relationship between sentence length and readability grade was larger (.58) than its relationship to whether the article had won an award (–.37), thus nullifying the relationship. A second path analysis that eliminated the path from sentence length to winner showed a significant path (b = –.25, p = .003) of the same size of the bivariate correlation between readability grade and winning articles. However, further analysis that compared these two models revealed that the

model in Figure 1 with the direct (but nonsignificant) path from readability to winner fits the data better than the one without this path ($\Delta\chi^2 = 12.64$, Δd.f. = 1, $p < .05$). It is noteworthy that the model in Figure 1 replicated across other measures of readability. Corresponding path analyses with the Flesch, Lix, or Kincaid scores as the definition of readability produced results with statistically significant paths of essentially the same size as when we used the average readability grade.

A possible counterexplanation for the observed differences between award-winning and nonwinning articles might be that numbers are counted as words of only a few characters and syllables. However, this cannot explain the greater readability of award winners for several reasons. First, word length did not differ between winners and nonwinners; only the length of sentences differed. Second, we included the article characteristics to control for such possible confounds. Third, all the *JMR* articles, winners and nonwinners, were empirical, but there were significant differences in readability. Thus, empirical status and the possibility of an artifactual result due to the way the readability software coded the word length of numbers cannot explain the difference between winners and nonwinners.

Readability scores for the four journals. A reason for the ANCOVAs was to test for differences among the journals in which the article was published. The results showed no meaningful differences among the four journals or presumed differences in prestige or impact. That there were no significant interactions between award winners and journal suggests that the results between award winners and nonwinners replicated across journals.

Characteristics of more readable articles. All the relationships specified in the paths in Figure 1 were statistically significant, except for the nil one between readability grade and whether the article had won an award. The two principal component scores used to summarize article types were significantly related to word length. As Figure 1 shows, there were shorter words in articles that reported experiments with behavioral conceptualization (b = –.21) and those that focused on methodological models analyzing secondary data (b = –.39). The only direct relationship of an article characteristic to sentence length was the number of references per sentence (b = +.18), which was also associated with worse readability grade levels (b = +.17). Thus, sentence length only partially mediated the effects of the number of references on readability. Longer sentences resulted in worse readability grades (b = +.58), as did longer words (+.52). These latter two results are near tautological because each of these two variables, in some form, is included in all the readability formulas.

Other results. Additional analyses showed that, consistent with Mick's (2005) logical conjecture, articles written by authors who have published more articles had better readability scores (r = –.27, $p < .001$). These experienced authors were also much more likely to write award-winning articles ($r_{\text{number of publications; award winner}} = .38$, $p < .001$). However, consistent with Hartley, Sotto, and Pennebaker's (2002) results and counter to the negative relationship that Stremersch, Verniers, and Verhoef (2007) find, the total number of an article's citations, the number of an article's citations per year, and coauthorship were not related to readability scores.

Discussion

Award-winning articles in AMA academic journals were more readable than nonwinners on several measures of readability. Readability, which includes measures of the length of words and sentences, was not directly influenced by any article characteristic, with one exception: the number of references. The number of references per sentence was associated with longer sentences, but sentence length did not completely mediate the effects of the number of references on readability. Thus, theoretical position and literature review articles, which often contain more references, were more likely to be rated as less readable beyond the merely tautological need for longer sentences to include the increased percentage of references. There were shorter words in articles that focused on behavioral experiments and methodological mathematical models that used secondary data than in other articles, such as articles on managerial topics, nonempirical articles, and articles that used survey data. However, word length completely mediated the effects of these various types of articles on readability.

Overall, these results show that with the possible exception of review articles, which may be less readable because of longer sentences, no particular type of article was more or less readable. The key is the author's discipline in writing clearly; the use of shorter words and shorter sentences can also help. According to Ross-Larson (1999), any sentence with more than 25 words is likely too long. Sentences of the articles in our sample averaged 25.1 words. In the classic *Elements of Style*, Strunk and White (2000, p. 23) succinctly advise to "omit needless words." Our analyses of the readability scores of journal articles support this advice. Sentence length was the key factor. Good sources for details about how to shorten sentences include Cook (1985), Flesch and Lass (1996), and Ross-Larson (1996, 1999). In addition to eliminating needless words, phrases, and sentences, Ross-Larson (1999, p. 18) suggests that writers "read a sentence aloud. If they stumble, gasp for air, the sentence is not well crafted, and the stumbles and gasps show them where to make repairs."

Brevity is a good thing, but it is not the prime goal of good writing. The key is improved reading ease and comprehension. Psychological research (e.g., Kintsch and Miller 1984) has shown that these factors are most highly correlated with readability, as measured by the formulas in the Appendix. Taken literally, winners and nonwinners differed by an average of less than a school grade. Although this difference may not be significant for academic researchers, DuBay's (2004, p. 49) review of readability research concludes that "even small differences in comprehension pay off." Small improvements in writing can help even expert readers read a manuscript faster, while understanding and remembering more.

Moreover, marketing is an applied discipline. Practitioners are more than half of the subscribers of AMA journals (73% for *JIM*, 68% for *JM*, 62% for *JMR*, and 53% for

TABLE 2
Differences Between Award-Winning and Nonwinning Articles

Measure	Award-Winning Articles	Nonwinners
Average readability grade	15.9	16.5
Kincaid grade level	14.2	14.9
Flesch index[a]	37.0	33.9
Sentence length (words)	24.0	26.5
Word length (characters)	5.3	5.4
Total number of words	9060	7203

[a]The Flesch score is the only readability index for which higher scores indicate better readability. It is not included in the mean readability grade scores.

Notes: All means are adjusted by three covariates in ANCOVA: the number of references per sentence in the article, whether it was more apt to be a behavioral experiment, and whether it was more apt to be a methodological modeling article that contained equations. The difference between winning and nonwinning articles for all but word length was statistically significant at $p < .006$. In their order in Table 2, effect sizes, as estimated by omega-squared, of the significant differences were .058, .063, .050, .158, and .195.

JPP&M). There is wide agreement that academics and practitioners do not communicate well (e.g., Crosier 2004). Several essays in the October 2005 issue of *JM* bemoan the lack of impact of academic research on marketing practice, and the writing style of many academic research articles may account for more than a little of this communication gap. If possible, authors should seek a review by a practitioner member of the target audience.

Although the statistically significant differences in the readability grade scores between award winners and nonwinners may seem small, the effect sizes (detailed in Table 2) approximated or exceeded what Cohen (1988; see also Sawyer and Ball 1981) defines for behavioral research as a medium size (6.25% explained variance). This difference is noteworthy because of the many other reasons readability scores might vary. Our comparison of winning and nonwinning published articles was a tough test of the discriminatory power of the readability indexes. There was a strong possibility of no significant differences among articles within top-level academic journals. We compared award winning articles with nonwinning articles, but the nonwinners were not "losers." They were published in AMA journals, which are selective and have high rejection rates.

Interpretation of the implications and importance of the practical effect size of the components of readability scores might be more enlightening than the scores themselves. Award-winning articles contained 2.5 fewer words per sentence (9.4%) than nonwinning articles. Given the 9900 words per award-winning article in *JM*, this 934-word advantage would amount to more than three manuscript pages and close to an average *JM* journal page—enough to accommodate an additional article in an issue that averages 10–11 articles per issue.[3]

[3]The technical editor at the AMA estimated that a typical manuscript page has slightly more than 300 words. Journal pages average approximately 930 words in *JMR*, approximately 1060 words in *JM* and *JPP&M*, and approximately 500 words in *JIM* (personal communication, 2005).

Conclusion

Motivation and perspiration are important ingredients for good writing. Although it is well known that multiple drafts are needed, many authors underestimate what is required for a well-written manuscript. Comments from colleagues and copy editors are imperative. Failing to maximize an article's readability is foolish and risky, leads to wasted time for journal reviewers, and increases the chances of rejection. Even after many drafts, authors should take a "final" draft and, one last time, try to simplify, shorten, and improve the structure of sentences and the manuscript as a whole. As Winer (1998, p. iv) advises, "consider the version you are thinking of sending as the n – 1st version. Give it to colleagues to read not only for content but for clarity as well."

We hope that our focus on good writing and the use of precise, short sentences to improve readability in marketing will motivate readers to work even harder in this regard. Although authors should not naively write simply to maximize readability scores, such scores are a useful screen—a "canary in the mineshaft" (Bauerly, Johnson, and Singh 2006, p. 223)—to predict reading ease.[4] Most manuscripts and sentences can be shortened and improved. As Holbrook (1986, p. 106) colorfully states, "Remember that brevity is the soul of wit. In each section, paragraph, and sentence of your paper, shun excessive length as diligently as you would avoid paying extra interest expenses on your credit card. Say exactly what you need to say, no more, then stop. Like this."

[4]After the elimination of the title, the abstract, the headings, the references, the tables, the figure, the footnotes, the quotations, and the Appendix, as was done for the articles in this research, this article had the following readability statistics: words per sentence = 23.8, characters per word = 5.3, readability grade = 14.8, Flesch readability level = 42.1, and Kincaid grade level = 13.3. For comparison, this readability grade score would have ranked about 32nd of the 162 articles in our analysis.

APPENDIX
Formulas for Readability Indexes

Kincaid grade level[a] = 11.8 × syllables/words + .39 × words/sentences – 15.59

ARI[a] = 4.71 × characters/words + .5 × words/sentences – 21.43

Fog grade level[a] = .4 × {words/sentence + 100 × [(words ≥ 3 syllables)/words]}

SMOG grade level[a] = square root of {[(words ≥ 3 syllables)/sentences] × 30} + 3

Lix[b] = words/sentence + 100 × [(words ≥ 6 characters)/words]

Flesch formula[b, c] = 206.835 – 84.6 × syllables/words – 1.015 × words/sentences

Average grade scale = (Kincaid + ARI + Fog + SMOG)/4

[a]One of four scores that are designed to provide a grade at which the material should be readable. Lower scores (grades) suggest more readable material.

[b]Scores varying between 1 and 100.

[c]Flesch is the only readability index for which higher scores indicate better scores.

REFERENCES

AMA (2005), "Evaluation Criteria for Manuscripts," (accessed on November 1, 2005), [available at http://www.marketingjournals.org/jm/ms_evaluate.php].

Armstrong, J. Scott (1980), "Unintelligible Management Research and Academic Prestige," *Interfaces*, 10 (April), 80–86.

——— (1989), "Readability and Prestige in Scientific Journals: Comment," *Journal of Information Science*, 15 (April), 123–24.

Bauerly, Ronald J. and Don T. Johnson (2005), "An Evaluation of Journals Used in Doctoral Marketing Programs," *Journal of the Academy of Marketing Science*, 33 (July), 313–29.

———, ———, and Mandeep Singh (2005), "Readability and the Impact of Marketing," in "Marketing Renaissance: Opportunities and Imperatives for Improving Marketing Thought," *Journal of Marketing*, 69 (October), 19–20.

———, ———, and ——— (2006), "Readability and Writing Well," *Marketing Management Journal*, 16 (1), 216–27.

Baumgartner, Hans and Rik Pieters (2003), "The Structural Influence of Marketing Journals: A Citation Analysis of the Discipline and Its Subareas Over Time," *Journal of Marketing*, 57 (April), 123–39.

Cohen, Jacob (1988), *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cook, Claire Kehrwald (1985), *Line by Line: How to Edit Your Own Writing*. Boston: Houghton Mifflin.

Crosier, Keith (2004), "How Effectively Do Marketing Journals Transfer Useful Learning from Scholars to Practitioners?" *Marketing Intelligence and Planning*, 22 (May), 540–56.

Dale, Edgar and Jeanne Chall (1949), "The Concept of Readability," *Elementary English*, 26, 19–26.

Desiraju, Ramarao and Steven M. Shugan (1999), "Strategic Service Pricing and Yield Management," *Journal of Marketing*, 63 (January), 44–56.

Dubay, William H. (2004), *The Principles of Readability*. Costa Mesa, CA: Impact Information, (accessed March 15, 2006), [available at http://www.nald.ca/fulltext/readab/readab.pdf].

Flesch, Rudolf and A.H. Lass (1996), *The Classic Guide to Better Writing: Step-by-Step Techniques and Exercises to Write Simply, Clearly and Correctly*. New York: HarperCollins.

Gencturk, Esra F. and Masaaki Kotabe (2001), "The Effect of Export Assistance Program Usage on Export Performance: A Contingency Explanation," *Journal of International Marketing*, 9 (2), 51–72.

Glazer, Rashi (1991), "Marketing in an Information-Intensive Environment: Strategic Implications of Knowledge as an Asset," *Journal of Marketing*, 55 (October), 1–19.

Golder, Peter N. and Gerard J. Tellis (1993), "Pioneer Advantage: Marketing Logic or Marketing Legend?" *Journal of Marketing Research*, 70 (May), 158–70.

Gopen, George D. (2004), *The Sense of Structure: Writing from the Readers Perspective*. New York: Pearson Longman.

——— and Judith A. Swan (1990), "The Science of Scientific Writing," *American Scientist*, 78 (November–December), 550–58.

Hartley, James, James Pennebaker, and Claire Fox (2003), "Using New Technology to Assess the Academic Writing Styles of Male and Female Pairs and Individuals," *Journal of Technical Writing and Communication*, 33 (3), 243–61.

———, Eric Sotto, and James Pennebaker (2002), "Style and Substance in Psychology: Are Influential Articles More Readable Than Less Influential Ones?" *Social Studies of Science*, 32 (April), 321–34.

———, Mark Trueman, and A.J. Meadows (1988), "Readability and Prestige in Scientific Journals," *Journal of Information Science*, 14 (April), 69–75.

Hoch, Stephen J., Xavier Dreze, and Mary E. Purk (1994), "EDLP, Hi-Lo, and Margin Arithmetic," *Journal of Marketing*, 58 (October), 16–27.

Hock, Roger R. (1992), *Forty Studies That Changed Psychology: Explorations into the History of Psychological Research*. Englewood Cliffs, NJ: Prentice Hall.

Holbrook, Morris B. (1986), "A Note on Sadomasochism in the Review Process: I Hate When That Happens," *Journal of Marketing*, 50 (July), 104–108.

Hult, G. Tomas M., William T. Neese, and R. Edward Bashaw (1997), "Faculty Perceptions of Marketing Journals," *Journal of Marketing Education*, 19 (Spring), 37–52.

Kintsch, Walter and James R. Miller (1984), "Readability: A View from Cognitive Psychology," in *Understanding Reading Comprehension*, James Flood, ed. Newark, DE: International Reading Association, 220–32.

Klare, George R. (1963), *The Measurement of Readability*. Ames: Iowa State University Press.

Knight, Jonathan (2003), "Scientific Literacy: Clear as Mud," *Nature*, 423 (May), 376–78.

Lutz, Richard J. (1990), "Editorial," *Journal of Consumer Research*, 17 (December), 242–45.

Medoff, Marshall H. (2003), "Article Placement and Market Signaling," *Applied Economics Letters*, 10 (June), 479–82.

Metoyer-Duran, Cheryl (1993), "The Readability of Published, Accepted, and Rejected Papers Appearing in *College and Research Libraries*," *College and Research Libraries*, 54 (November), 517–26.

Mick, David (2005), "Inklings: From Mind to Page in Consumer Research," *Association for Consumer Research Newsletter*, 1 (Summer), 1–3.

Morgan, Neil A., Anna Kaleka, and Constantine S. Katsikeas (2004), "Antecedents of Export Venture Performance: A Theoretical Model and Empirical Assessment," *Journal of Marketing*, 68 (January), 90–108.

Muthén, Linda K. and Bengt O. Muthén (2004), *Mplus User's Guide*. Los Angeles: Muthén & Muthén.

Oppenheimer, Daniel M. (2005), "Consequences of Erudite Vernacular Utilized Irrespective of Necessity: Problems with Using Long Words Needlessly," *Applied Cognitive Psychology*, 20 (October), 139–56.

Peter, J. Paul and Jerry C. Olson (1983), "Is Science Marketing?" *Journal of Marketing*, 47 (Fall), 111–25.

Prasad, V. Kanti, Keshavamurthy Ramamurthy, and G.M. Naidu (2001), "The Influence of Internet-Marketing Integration on Marketing Competencies and Export Performance," *Journal of International Marketing*, 9 (4), 82–110.

Redish, Janice C. and Jack Selzer (1985), "The Place of Readability Formulas in Technical Communication," *Technical Communication*, 32 (August), 46–52.

Robinson, William T. and Sungwook Min (2002), "Is the First to Market the First to Fail? Empirical Evidence for Industrial Goods Businesses," *Journal of Marketing Research*, 39 (February), 120–28.

Ross-Larson, Bruce (1996), *Edit Yourself*. New York: Norton.

——— (1999), *Stunning Sentences*. New York: Norton.

Sawyer, Alan G. (1988), "How to Write a Painful Marketing Manuscript: Prose and Cons," *Journal of Marketing Education*, 10 (Fall), 49–53.

——— and A. Dwayne Ball (1981), "Statistical Power and Effect Size in Marketing Research," *Journal of Marketing Research*, l8 (August), 275–90.

Stewart, David W. (2002), "Getting Published: Reflections of an Old Editor," *Journal of Marketing*, 66 (October), 1–6.

Stremersch, Stefan and Gerard J. Tellis (2002), "Strategic Bundling of Products and Prices: A New Synthesis for Marketing," *Journal of Marketing*, 66 (January), 55–72.

———, Isabel Verniers, and Peter C. Verhoef (2007), "The Quest for Citations: Drivers of Article Impact," *Journal of Marketing*, 71 (July), 171–94.

Strunk, William, Jr., and Elwyn B. White (2000), *The Elements of Style*. Needham Heights, MA: Allyn and Bacon.

Taylor, David and Intuitive Systems (2006), "Readability.info," (accessed May 7, 2006), [available at http://www.readability.info/uploadfile.shtml].

Winer, Russell S. (1998), "From the Editor," *Journal of Marketing Research*, 35 (February), iii–v.

Vorhies, Douglas W. and Neil A. Morgan (2003), "A Configuration Theory Assessment of Marketing Organization Fit with Business Strategy and Its Relationship with Marketing Performance," *Journal of Marketing*, 67 (January), 100–115.

Zinsser, William (1988), *On Writing Well: An Informal Guide to Writing Nonfiction*. New York: Harper and Row.