How does the peer review process influence AANA Journal article readability?

CHUCK BIDDLE, CRNA, PhD
Lebanon, New Hampshire
JOHN AKER, CRNA, MS
Iowa City, Iowa

This study examined the readability of the AANA Journal, quantifying the effect of peer review on case and research reports published from 1992 to 1994. Gunning and Flesch index-based computer analysis, as well as human comparative analysis, was undertaken. Computer and human assessment of readability revealed improvement as papers evolved from submitted to published versions; however, at publication the manuscripts remained in the “difficult” readability range. Although this study provides evidence that peer review improves readability, it may be that, due to a professed need for scientific purity and an imposed sense of scholarship, nursing and other biomedical journals may over-emphasize a style and approach that paradoxically make transfer of information unreasonably difficult.

Key words: Nursing publications, peer review, readability of manuscripts.

A great deal of adult learning is accomplished by way of written communication. Journal publications represent an efficient and time-honored manner for disseminating practical and theoretical information. In a sense “one is what one reads.” There is substantial evidence that the abilities to critically appraise and understand professional literature are crucial components of being a quality provider.¹

Nurses rely on the written word for clinical and professional information. Readers presume that because the material they encounter in professional publications is peer reviewed, it is accurate, up-to-date, and edited in a fashion to make it consumable. Yet little systematic study has been directed toward the successes and failure of peer review.

While the readability of manuscripts in clinical journals should be high, evidence shows that it is not. For example, subscribers to the peer-reviewed Annals of Internal Medicine complained that it was too hard to read; only 1% indicated that they read the journal through to the end.² Furthermore, the process of peer review, while generally regarded as indispensable, has been noted to be flawed in a number of ways.³ Common criticisms of peer review include the use of unqualified or biased reviewers, absence of outcomes suggesting that reviewers “make a difference,” and failure to ensure the clinical relevance of articles.

Despite a search of the literature and queries to several publishers of healthcare journals, no systematic observations of the readability or user friendliness of nursing journal articles were uncovered. One purpose of a nursing journal is to promote the development of its readers’ practice; as such, the readability of the journal should be high. The purpose of this study was to examine the readability of an advanced practice nursing
Methods

The AANA Journal is published bimonthly and is probably examined by some 40,000 nurses, physicians, pharmacologists, and scientists. It is a clinically based journal targeting the advanced practice nurse in anesthesia. It publishes letters, case reports, book reviews, editorials, product evaluations, review articles, and original research. From the time of receipt until final publication, each manuscript is evaluated by a minimum of three advanced practice nurses or physician reviewers, a nurse editor-in-chief, a nonnurse associate editor, a nonnurse publications manager, and the author again at the galley proof stage. The acceptance rate for unsolicited manuscripts averages about 11% with virtually all requiring one major and one minor revision before publication.

Published case reports and reports of original research submitted from January 1992 to December 1994 were studied. Manuscripts were evaluated in their original form, then reevaluated. Each manuscript was placed on computer disk and retrieved to undertake the analysis at both the original (as submitted) and the final (as published) versions. Only the main texts of the articles were studied; abstracts, figures, legends, appendices, and the like were not studied because their peculiar organization would have hampered computer assessment. All documents were analyzed collectively with no attempt to identify differences between case and research reports.

Computer analysis of all manuscripts was accomplished using a program that incorporated the Gunning and Flesch indexes (Microsoft Word 6.0 on a Macintosh Centris 610 computer). Flesch's classic article introduced a well-accepted and proven reading ease index amenable to computer analysis. His scale generates a reading ease score from 0 to 100, with 100 being easiest. A second index, the Gunning Fog Index, similarly evaluates the degree of difficulty of a document and weighs the reading difficulty in terms of the number of years of education required to achieve readability and comprehension. This formula was set forth in Gunning's seminal book, *The Technique of Clear Writing.* Table I provides a brief overview of the relative weighing of the scales.

<table>
<thead>
<tr>
<th>Scalea</th>
<th>Flesch Scaleb</th>
<th>Interpretation</th>
<th>Illustration</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-29</td>
<td>&gt;16</td>
<td>Extremely difficult</td>
<td>Legal contract</td>
</tr>
<tr>
<td>30-49</td>
<td>14-16</td>
<td>Difficult</td>
<td>Ulysses by James Joyce</td>
</tr>
<tr>
<td>50-59</td>
<td>12-13</td>
<td>Somewhat difficult</td>
<td>The Atlantic Monthly</td>
</tr>
<tr>
<td>60-69</td>
<td>9-11</td>
<td>Average</td>
<td>The New York Times</td>
</tr>
<tr>
<td>70-100</td>
<td>5-8</td>
<td>Easy</td>
<td>Reader's Digest</td>
</tr>
</tbody>
</table>

a. Reading ease score  
b. Reading grade level

Published reliability and validity coefficients for these instruments range from 0.62 to 0.92 over a range of lay and scientific publications, although coefficients specific to the disciplines of nursing and medicine are not available. To address this issue, two doctoral-prepared nonanesthesia nurses rated 10 different research reports randomly sampled from the pool that underwent computer analysis. Half were in the submitted form and half in the published form; raters were unaware of the manuscripts' status and rated each using the Flesch and Gunning indexes. During the initial training sessions, these nurses demonstrated strong intrarater reliability with coefficients of agreement ranging from 0.79 to 0.88 (at the 0.05 to 0.01 confidence level) on six documents published in diverse sources, namely Reader's Digest, *The New Yorker, The Boston Globe, Nursing Research, Heart & Lung,* and The New England Journal of Medicine.

A one-tailed paired t test was used to compare submitted and published versions in terms of length and readability. (A one-tailed test was selected as we hypothesized that readability would improve.) A two-tailed t test was used to compare human and the computer analyses of readability. A prestudy alpha of $P < 0.01$ was established for all comparisons.

Results

Fifty-nine manuscripts constituted the sample; of these 26 were case reports and 33 were reports of original research. The document statistics presented here are reported as a mean followed by the
standard deviation. The length of case reports at the submission stage was 2,793 ± 973 words and at the publication stage it was 2,371 ± 840 words. The length of research reports at the submission stage was 4,842 ± 1,225 words and at the publication stage it was 3,609 ± 1,043 words. The number of words per sentence for case reports was 25.8 ± 2.1 at the submission stage and 18.9 ± 4.6 at the publication stage. The number of words per sentence for research reports was 24.6 ± 5.0 at the submission stage and 19.1 ± 3.2 at the publication stage. All reported group differences between the submission and publication stages were significant (P < 0.01).

**Computer and human analysis of readability.**

Table II summarizes the results of the computer and nurse raters' analysis of readability. A significant improvement in readability (P < 0.01) was demonstrated with both the Flesch and Gunning indexes irrespective of whether the analysis was performed by computer or the humans. There was no significant difference (P > 0.01 by t test) between the computer and human analyses of readability.

**Discussion**

The present study indicates that a significant improvement in readability resulted as manuscripts navigated the peer review and editorial process. Contraction of manuscript size and improvement in reading ease was observed with both case and research reports. This suggests that peer review accomplished a positive aim, improving (making easier) readability. While this appears laudatory, it should be noted that at publication the manuscripts remained in the "difficult" readability range. While the match of reading "grade level" and audience seems reasonable, perhaps readers' needs would be better served by making the published reports even more "readable."

That no significant differences were found between the computer and human assessments of readability suggests that the Flesch and Gunning indexes provide useful and valid measures of readability of clinical and research nursing literature. Further, examination of the reliability and validity of these tools in other biomedical literature is recommended.

Peer review and editor modification are foundations by which the art and science of nursing are communicated to journal readers; this study must be repeated with other nursing journals before changes in these traditions can be recommended. However, it is tempting to urge some consideration toward making our publications more user-friendly. Perhaps due to a professed need for scientific purity and sense of scholarship, nursing and biomedical journals have overemphasized a style and approach that paradoxically make transfer of information difficult if not daunting.

While recognizing that nursing journals have audiences ranging from the bench researcher to the bedside nurse, editors must be mindful that nursing is a practice-based profession grounded fundamentally in managing patient-care problems. As such, the ultimate consumer of a published article is the clinician caring for the patient. While it is essential for nurses to generate and publish scholarly papers, researchers, writers, and editors must strive to "keep their audience in mind" as ideas and findings are developed and disseminated. It is hoped that this study will stimulate further systematic inquiry directed toward the issues of peer review and readability of nursing publications.

### Table II
**Evaluation of readability**

<table>
<thead>
<tr>
<th>At initial submission</th>
<th>At publication</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Case reports</td>
</tr>
<tr>
<td><strong>Computer analysis</strong></td>
<td>n = 26</td>
</tr>
<tr>
<td>Flesch Reading Ease</td>
<td>27.14 ± 8.60</td>
</tr>
<tr>
<td>Gunning Fog Index</td>
<td>18.20 ± 3.11</td>
</tr>
<tr>
<td><strong>Human analysis</strong></td>
<td>n = 10</td>
</tr>
<tr>
<td>Flesch Reading Ease</td>
<td>26.92 ± 5.16</td>
</tr>
<tr>
<td>Gunning Fog Index</td>
<td>18.23 ± 6.47</td>
</tr>
</tbody>
</table>

All values reported as mean ± standard deviation
All pre- and at-publication comparisons were significant (P < 0.01 by t test)
All comparisons between computer and humans were not significant (P > 0.01 by t test)
REFERENCES


AUTHORS

Chuck J. Biddle, CRNA, PhD, is a graduate of the Norfolk, Virginia, General Hospital School of Anesthesia. His bachelor's and master's degrees were earned at Old Dominion University there and his doctoral degree from the University of Missouri, Kansas City, Kansas. He is chief nurse anesthetist in the Department of Anesthesiology at Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire. Dr. Biddle serves as editor in chief of the AANA Journal.

John G. Aker, CRNA, MS, received his BS degree from the University of Topeka and his MS degree in anesthesia from the University of Kansas Medical Center, Kansas City, Kansas. He is a staff nurse anesthetist at the University of Iowa Hospitals and Clinics in Iowa City. He served on the AANA Journal Editorial Advisory Board and is a member of the International Anesthesia Research Society.

This paper underwent blinded review that was under the supervision of an appointed guest editor.