5-1-1993

Quality of Abstracts

Carol Tenopir
University of Tennessee - Knoxville

Peter Jasco

Follow this and additional works at: https://trace.tennessee.edu/utk_infosciepubs

Part of the Library and Information Science Commons

Recommended Citation
Tenopir, Carol and Jasco, Peter, "Quality of Abstracts" (1993). School of Information Sciences -- Faculty Publications and Other Works. https://trace.tennessee.edu/utk_infosciepubs/128

This Article is brought to you for free and open access by the School of Information Sciences at TRACE: Tennessee Research and Creative Exchange. It has been accepted for inclusion in School of Information Sciences -- Faculty Publications and Other Works by an authorized administrator of TRACE: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.
Quality of abstracts

Authors: Carol Tenopir and Peter Jacso
Date: May 1, 1993
From: Online (Vol. 17, Issue 3.)
Publisher: Information Today, Inc.
Document Type: Article
Length: 4,800 words

Abstract:
Abstracts allow on-line searchers to view a short version of an original document. Abstracts also provide additional search terms to help focus a review of an on-line, printed or CD-ROM database. An abstract may be written by the author of an original article; however, professional services typically write their own abstracts. Abstracters may be hired for writing skills, but in some cases abstracting and indexing tasks may be combined. The quality of abstracts can be judged by readability and informativeness.

Full Text:
Given a choice between an index without abstracts and one with abstracts, we suspect most users would pick the abstract file every time. Abstracts help users judge the relevance of articles on a subject, they provide a handy synopsis of an article's contents, and sometimes, they serve as a substitute for the original document. On most online systems and CD-ROM versions, abstracts provide additional searchable words as well.

Printed, online, and CD-ROM indexing and abstracting publications are familiar reference tools in libraries. Many, such as ERIC, Psychological Abstracts, or Chemical Abstracts, are mainstays of libraries' collections in one or more formats because they cover their topic areas with a depth or quality unmatched by any competition.

In other subjects, librarians must make more difficult selection decisions. Several indexes may cover the same topic or include many of the same sources. One feature that may be a factor in the evaluation and selection process is the presence of abstracts or the quality of the abstracts.

Competition is especially tough among several similar periodical indexes on CD-ROM. H.W. Wilson Company's Readers' Guide Abstracts (RGA), UMI's Periodical Abstracts OnDisc (PAO) (and its subfile, Resource-One), and EBSCO's Magazine Article Summaries (MAS) all index and abstract general-interest magazines for the layperson [1]. In 1989 Tenopir and Smith provided an in-depth comparison of these general interest products (plus indexing-only products from Information Access Company) [2]. One factor that differentiates these indexes and could be an important consideration in evaluation is the quality of their abstracts.

WHAT IS AN ABSTRACT?

To judge the quality of abstracts, it is necessary to understand what makes a good abstract. Judgements of "good" are always dangerous, but, in the case of abstracts, there is some consensus among the experts. The American National Standards Institute (ANSI) developed a standard for abstracts in 1979 (it is currently being reviewed for reaffirmation) that in many ways dictates what a good abstract should be.

ANSI defines an abstract an "an abbreviated, accurate representation of the contents of a document" [3]. Robert Collison, in a book devoted to abstracting, describes an abstract as:

...the terse presentation in (as far as possible) the author's own language, of all the points made, in the same order as in the original piece of primary documentary information....It is much more than this, too, for it is intended to stand as a readable and complete item in its own right, and is therefore a separate work of scholarship that can be fully indexed and exploited for the benefit of scholars and research workers [4].

WHO WRITES ABSTRACTS?

According to ANSI it doesn't matter who writes an abstract--it may be written by the author of the original work or by a professional abstractor--as long as it follows certain quality guidelines. The three general periodical indexes studied here (Readers' Guide Abstracts, Periodical Abstracts OnDisc and Magazine Article Summaries) use staff abstractors. On the other hand many indexes of
An informative abstract is a condensed version of the essential ideas of a document. It contains a statement, from the viewpoint of the author, of the thesis, development or proof and conclusions. In short, it states what the document actually says [13].

TYPES OF ABSTRACTS

Good abstracts do not all look exactly the same. ANSI recognizes three legitimate types of abstracts (indicative, informative, and combination), while some sources recognize up to five types (indicative, informative, combination, critical, and special purpose) [11]. Figure 1 shows examples of each major type of abstract.

Indicative (descriptive) abstracts tell readers what they will find if they go to the complete article. They describe the kinds of information in an article, but they do not reproduce that information. Indicative abstracts often contain such words as "includes, discusses, describes, are given" and the like, but rarely contain actual findings. They help users judge the likely relevance of the original article and provide additional online access points, but they cannot substitute for the original.

The ERIC Processing manual: Abstracts describes in detail the contents and purposes of an indicative abstract:

An indicative abstract is a description of or guide to the content and format of the document, written from the viewpoint of an informed but impartial and objective reader. The indicative abstract reports broadly what is discussed or included in the document, in what manner the information is presented, and if necessary, to whom the document is addressed [12].

Indexing and abstracting services may choose to produce indicative abstracts because they are easier and much faster to write than other types. In fact, according to ERIC instructions, "the abstractor may obtain sufficient information to write an indicative abstract by examining the Table of Contents, Foreword, Introduction, Summary, etc. or by scanning the text." Indicative abstracts may be necessary for certain kinds of original documents such as bibliographies, conference proceedings, review articles, and other documents where the content varies widely.

Informative abstracts are the preferred type by both ANSI and ERIC. Informative abstracts may sometimes substitute for the original document because they contain the actual findings or viewpoints of the original. According to the ERIC Processing Manual:

An informative abstract is a condensed version of the essential ideas of a document. It contains a statement, from the viewpoint of the author, of the thesis, development or proof and conclusions. In short, it states what the document actually says [13].

Writing a good informative abstract may be more difficult because it requires reading all (or most) of the original document and concisely summarizing its main points. It becomes an example of Collison's "readable and complete item in its own right," but contains only the original author's views, never the voice of the abstractor.
Some types of articles don't lend themselves to either indicative or informative abstracts, so ANSI allows a combination informative-indicative type. Usually combination abstracts are used for lengthy documents that cannot be completely covered informatively. The most important parts are abstracted in the informative style, while the abstractor "relegate[s] other aspects to indicative statements" [14]. Documents with many tables, graphs, charts, and appendices are best handled with combination abstracts.

Other types of abstracts are not recognized as valid by ANSI, but still are used by some abstracting services. Critical abstracts make a value judgment of the contents of an article where the abstractor deliberately includes his or her opinions of the merit of the article being abstracted. Most indexing and abstracting services warn against critical abstracts. Cleveland and Cleveland question whether they truly constitute a "third category." "If it is heavily editorial, it cannot convey much basic information and is really a review of the document rather than a true indicator of document content" [15]. According to the Clevelands, criticism should be reserved for reviews.

Special purpose abstracts are designed to meet the specific objectives of an abstracting service or an organization [16]. They reflect the point-of-view of the intended audience or include only information from the part of an article that is of special interest to a selected audience.

For example, abstracts of articles from the Journal of the American Medical Association (JAMA) that are in Health Periodicals Database are written especially for the layperson. The abstractor's intent is to be understandable and to stress the sections of the articles that have immediate practical application. Abstracts of these same articles may be quite different in MEDLINE, where the intended audience is health professionals.

If the ANSI authority is accepted, one general test of quality then is the type of abstracts in a database. Informative abstracts should predominate, along with some indicative or combination informative-indicative abstracts in some special cases. General purpose abstracting tools, aimed at a mixed, general audience should never have critical or special purpose abstracts.

WHAT GOES INTO AN ABSTRACT

The procedures for writing abstracts may have some bearing on the ultimate quality. According to Cleveland and Cleveland, an abstractor reads or scans a document, looking for five overall indicators of content [17]:

1) objectives and scope 2) methodology 3) results or outcomes 4) conclusions 5) any other supportive or peripheral information

Clearly by these procedures, a well-structured research article is easier to abstract than an opinion piece or most general interest magazine articles.

ERIC instructs its abstractors to consider eight things when reading a document to abstract [18]: 1) subject matter, scope, purpose 2) publication/document type 3) author's viewpoint 4) intended audience 5) relationship of this work to other works 6) intended use 7) special features such as glossary, maps, etc. 8) results or findings

Abstracts do not include: historical background, introduction, summaries, redundancy, old information, details of standard procedures, information the reader is expected to know, ideas for future research, raw data, every finding, or personal prejudices or opinions of the abstractor [19].

One "major facet of quality" of abstracts, according to Lancaster, is whether the major "points" of the document are brought out in the abstract [20]. This facet is directly influenced by careful application of the above guidelines or similar procedure to elicit information for the abstract. Information left out of the abstract is lost to the user of the indexing and abstracting publication.

LENGTH OF ABSTRACTS

The length of an abstract is not necessarily an indicator of quality. ANSI states that "[f]or most papers and portions of monographs, an abstract of fewer than 250 words will be adequate." Borko and Bernier recommend a length of between one tenth to one twentieth of the original for science literature [21]. However in reality, the desirable length of an abstract is a function of the length of the original articles and the policy of the abstracting service.

Lancaster suggests:

it makes sense that the length should vary with such factors as the length of the item itself, its range of subject matter, its perceived importance, its physical availability, and its intellectual accessibility (e.g., items difficult to locate, such as conference papers, or items in obscure languages might be abstracted in more detail than other items) [22].

On the other hand, consistently providing very short abstracts may be a disservice to the user. Without an average of at least 100 words, abstracts do little more than tantalize readers, rather than inform them. Average lengths of between 100 and 250 words appear ideal.

WRITING STYLE

Another important measure of quality must be the clarity and readability of the final product. Lancaster's second facet of quality is "are these points described accurately, succinctly, and unambiguously?" [23]
ANSI provides guidelines for writing style, including required use of paragraphs and complete sentences; preferred use of a third-person point of view and active over passive voice; sparing use of nontextual material (use only when necessary); and avoidance of unfamiliar terms, acronyms, abbreviations, or symbols [24].

COMPREHENSIVE CRITERIA FOR ABSTRACTS

Lancaster describes many different attempts to define the quality of abstracts and he gives examples of several quality criteria lists developed by individual indexing companies or other organizations [25]. The most useful list of criteria and methods for judging the quality of abstracts is still the list developed in 1975 by Borko and Bernier [26].

According to them, quality of abstracts can be judged in several ways:

1. A global rating of quality (by human judges)
2. The extent to which the ANSI or another standard is observed
3. The inclusion of significant information and the exclusion of unimportant information
4. Lack of errors
5. Consistency of style and readability
6. Relevance predictability
7. Ability to serve as a surrogate for the original (informative abstracts)
8. Adequacy as a source of index terms

Some of these criteria are clearly difficult to measure. The inclusion of significant information (#3) and lack of errors (#4), for example, are highly subjective. No two abstractors will write an abstract exactly the same or even agree about what information is to be included.

TESTING THE QUALITY OF ABSTRACTS IN THREE DATABASES

We decided to test the most measurable of these criteria to compare the quality of the abstracts found in the three major CD-ROM general periodical indexes that contain abstracts. Using quantitative or objective tests removes the potential variation caused by different viewpoints and provides a generalizable view of quality.

We chose to look at:

1) consistency of style and readability
2) the extent to which the ANSI standard is observed
3) informativeness

READABILITY OF ABSTRACTS

Readability of any text is a combination of factors that make a text comprehensible. It is a relative measure, usually based on grade level or age at which a text can be easily comprehended. Several formulae of readability have been widely accepted for testing the reading or grade level of texts. The most commonly used readability tests are the Flesch Reading Ease Score, Gunning's Fog Index, and the Flesch-Kincaid Grade Level. Each applies a formula to texts that purportedly measures how easy or difficult the text is to read. Each uses a variation of counting the average number of syllables per word, words per sentence, or percentage of words of three syllables or more [27].

Manually calculating indexes of style and readability have the serious disadvantage that they require lots of time. Fortunately all three tests are widely available in style and grammar software, such as Grammatik, used for analyzing machine-readable texts. This style analysis software allows readability analyses to be done without much effort on records extracted from online or CD-ROM databases.

We selected records to test from the database by author, subject, journal title, publication period, document type, original language, etc. These criteria can be used to adjust the composition and size of the sample to be evaluated.

Downloading a sizable sample may take some time (and money in the online environment), but it is the computer's time and not yours. If the software allows you to download only the abstract field you may immediately feed your file of abstracts to the style analyzer software. Most online systems and the popular CD-ROM software allow the user to specify the fields to be downloaded.

If the software does not allow downloading only the abstract field, then you must write a macro in your word processing software to strip the irrelevant fields from the downloaded records (accession number, title, author, descriptors, etc.). You must examine the structure of the records of each database to recognize the pattern, and to come up with an unambiguous algorithm. You also must
This kind of stripping may be needed even if the software allows the downloading of the abstract field alone, because the abstract may include data elements that will distort the result of the analysis if retained. For example, the abstracts in PsycLIT include the notice "Copyright 1992 American Psychological Assn., all rights reserved." Your macro should be programmed to delete such sentences. The stripped file can then be fed to a style analyzer to do the readability analysis.

Rightwriter and Grammatik are two widely used standalone style analyzers. Microsoft Word has a built-in style analyzer. All three provide paragraph, sentence and word statistics, as well as the Flesch Reading Ease Score, Gunning's Fog Index, and the Flesch-Kinkaid Grade Level measure. For this article Grammatik-IV was used.

READABILITY TESTING: SAMPLE RESULTS

Figure 2 shows the comparative analysis of 1,070 abstracts downloaded from the Resource-One database, 999 abstracts from Readers' Guide Abstracts (RGA), and 984 abstracts from Magazine Articles Summaries (MAS). For this analysis, records were retrieved from each database by a subject-oriented query, searching for articles about airlines. The Flesch-Kinkaid Grade Level and Gunning's Fog Index provide scores that correspond to standard American grade levels. (Grade 12 is senior in high school, while Grade 17 is one year of graduate school.)

The Flesch Reading Ease Score may range from 0 to 100. The lower the score the more difficult the text is to read. A score between 0-30 means very difficult, between 30-50 difficult, 50-60 fairly difficult, 60-70 standard. Figures above 70 indicate writing that is fairly easy to very easy-to-read.

A high percent of passive voice and prepositions, too many sentences per paragraph, too many words per sentence, and too many syllables per word make the reading more difficult. While the style analyzer program interprets the scores for you, they make more sense when you see the scores relative to the scores of other well-known writing. Grammatik presents a series of charts that compare the text analyzed with the scores of three other writings: the Gettysburg Address, a Hemingway short story, and a life insurance policy.

According to the scores the abstracts of Resource-One are the easiest to read. It has the highest Flesch Reading Ease score, and is two grade levels lower than MAS and RGA. Note, on the other hand, the high percentage of passive voice in Resource-One, which is known to make the reading of a text more difficult. Though the style analyzers are fast and consistent in calculating scores and indexes, they are not infallible, and your judgment may be required to adjust the results.

For instance, the sentence length score is not necessarily a positive feature of Resource-One records. The reason for this score is that the last sentence of the abstract is often nothing more than an indication of the type of illustration available, usually a single word or two, such as "Tables," or "Graph and Chart." Similarly, the abstracts of Magazine Article Summaries records typically include the author name as the last sentence, with the first name only as an initial, lowering the average sentence length significantly. Another reason for the fairly low average of sentence length is the many terse abstracts in MAS, not unlike the TV-program synopsis and notes (Boy meets, loves, leaves girl; language, nudity, violence).

The length of the paragraphs give Readers' Guide Abstracts the lowest scoring in readability, but also this may be one of its biggest assets. (The style analyzers consider that one abstract is one paragraph in the downloaded file, and the Readers' Guide abstracts are two and a half times as long as the others as shown by the analyzer.)

(Just for the record, Grammatik gave this article a Flesch-Kincaid grade level of 12, a Flesch reading ease score of 42, and a Gunning's Fog Index of 15. The usage of passive voice was only 9%, and the average sentence length was 16.1 words. Average paragraph length was 2.7 sentences.

TESTING READABILITY ANOTHER WAY

Another way to judge abstracts is to select records by journal title to see how abstracts from two or more databases compare when the sources are identical. Figure 3 shows the result of the analysis of over 300 abstracts of PsycLIT and sociofile for articles published in the Journal of Social Psychology. (sociofile abstracts include the indication of references, illustrations, and the source of the abstract, all of which affect the statistics.)

Not surprisingly, the abstracts of these databases indicate a high grade level, feature longer words, and a higher use of passive voice than the general interest databases. The ease of reading score for PsycLIT is better than for sociofile, but the average length of abstracts in the latter seems to be far longer than in PsycLIT. Again, you must consider database peculiarities. It is possible that one database contains information in the abstract that is in separate fields in the other database.

The above technique can be applied, of course, not only to abstracts, but also to the ever-growing number of full-text databases. You can analyze, for example, the editorials in U.S. News and World Report, or the movie reviews in New York Times versus the Washington Post. To put the results of full-text analysis in perspective it is interesting to study the Flesch-Kincaid grade level indexes of some widely known publications as reported in the April, 1991 issue of PC Sources.

COMPATIBILITY WITH ANSI

Another measure of quality and compatibility with the ANSI standard, is also for the most part an objective measure, but requires
more human labor and judgment.

One part of the question is easy to answer from word processing or style analyzer software. ANSI recommends abstracts of between 100 to 250 words. Figure 4 shows average lengths for the five databases tested. Abstracts for scholarly journal articles aimed at a professional audience might be expected to be longer than those for general periodical articles.

Note, however, that Readers’ Guide has abstracts that average 110 words, almost twice as long as Magazine Article Summaries and three times as long as Resource-One. By this criteria, then, Readers’ Guide Abstracts most closely matches the ANSI standard for recommended length. Even though all three general interest databases index and abstract almost the same magazines (and, therefore, the same length of articles) the average length of abstracts among the three varies considerably.

INFORMATIVENESS

Another part of the standard, the recommendation of informative abstracts, cannot be so easily tested. This requires reading a sample of records from each database and judging whether the abstract is informative, indicative, a combination of informative/indicative, or another type. To do this you must be familiar with the ANSI standard and types of abstracts.

We examined 300 randomly selected records for each database from the subset of records that we downloaded for the readability measures. Figure 5 compares type of abstracts for the three general-interest databases. None of the databases contained any critical or special purpose abstracts. The percent of informative, indicative, or combination varied, however.

Informativeness of a database can be measured by the percent of informative abstracts included. Second best is a combination informative/indicative. If a database has a high percentage of indicative abstracts, it not only does not meet the ANSI standard, it rates low on informativeness. Users of this database probably always have to turn to the original source, instead of relying on the abstracts.

Again, Readers’ Guide Abstracts best meets this quality test. Over 81% (245 of 300) are informative. Only 9.7% (29) are indicative alone. Magazine Article Summaries and Resource-One rely on indicative and combination abstracts.

Consider just one word of caution. The Magazine Article Summaries indicative abstracts may in some ways be better than the Resource-One informative abstracts because the short length can mislead the reader. Magazine Article Summaries abstracts are mostly indicative or combination. They tell the reader what is in the article (and therefore what is not in the abstract.) They are very brief, but the reader will know he or she must go to the original to find the information briefly described in the abstract. Resource-One's combination or "informative" abstracts usually give just one or two facts from the article, and fail to indicate the full range of information that could be found by going to the original.

SUGGESTIONS FOR FUTURE STUDY

This study looked at objective global measures of quality of abstracts. Other aspects of quality are relative and are subject to interpretation. What is acceptable and appropriate in some libraries for some library users may not be so for others. This especially struck us when we heard of a community college instructor and an intermediate school librarian who preferred an index with shorter, less informative abstracts because they didn't want students to rely on abstracts without having to go to the original article.

Another judgment of quality may thus be the judgment of how appropriate the abstracts are for a given constituency. A future study might ask librarians from schools, community colleges, small four-year colleges, universities, and public libraries, along with end-users from a variety of environments, to judge the relative quality and appropriateness of abstracts. It would be interesting to learn if their judgments differed and to discover how their judgments compare with the more objective measurements discussed in this article.

REFERENCES

[1] To be fair, EBSCO's Magazine Article Summaries do not claim to be full abstracts. By calling them "summaries" perhaps EBSCO intends to avoid the kind of in-depth analysis and comparison we are doing. However, since they are used in libraries in the same way as "abstracts," we think such a comparison is justified.


[17] Cleveland and Cleveland. Introduction to Indexing and Abstracting.


[19] Ibid.


[22] Ibid., p. 101.

[23] Ibid., p. 105.


Copyright: COPYRIGHT 1993 Information Today, Inc.
http://www.infotoday.com/default.asp

Source Citation (MLA 8th Edition)

Gale Document Number: GALE|A13828066