ing CA Index Names and their CAS Registry Numbers, it serves as a somewhat limited printed version of the Name Match system. To ensure the accuracy of the Index Guide data base, the file has been checked against the Registry Nomenclature File. Because CAS uses a systematic chemical name whenever the chemical structure is known, the Index Guide cross-references may also be used to identify names or retrieve Registry Numbers.

In summary, the CAS Registry Name Match System provides a rapid and efficient way of equating by machine the many commercial product names encountered in a wide variety of technical literature with the systematic chemical substance index names and Registry Numbers to be used in the CAS products. In a similar way, the CA Index Guide provides to the user of CAS products a manual way to match many commercial product names with CA Index Names and Registry Numbers.

LITERATURE CITED

(1) CA Volume 76 Index Guide (1972), Section IV; Donaldson, N., Powell, W. H., Rowlett, Jr., R. J., White, R. W., and Yorka, K. V., "CHEMICAL ABSTRACTS Index Names for Chemical Substances in the Ninth Collective Period (1972–1976)," *J. Chem. Doc.*, **14**, 3–15 (1974).

(2) Morgan, H. L., "The Generation of a Unique Machine Description for Chemical Structures—A Technique Developed at Chemical Abstracts Service," *J. Chem. Doc.*, **5**, 107–113 (1965); Rowlett, Jr., R. J., and Tate, F. A., "A Computer-Based System for Handling Chemical Nomenclature and Structural Representations," *J. Chem. Doc.*, **12**, 125–128 (1972).

(3) CA Volume 76 Index Guide (1972), Section II, paragraph 10, p 71.

A Study of Citations to 308 Journal Articles in Chemistry Published in 1963

RICHARD W. GREFRATH

Mortvedt Library, Pacific Lutheran University, Tacoma, Washington 98447

Received March 26, 1974

A study was made of the behavior of citations to 308 journal articles in the field of chemistry published in 1963, and the citations to the articles were tabulated year by year to the present time (through the 1971 cumulation). The articles were classified as being practical or theoretical and by authors in college/university or in industry settings. The hypothesis was that theoretical articles would have a longer life than practical articles. Combinations of the various classifications of articles were made (*e.g.*, theoretical articles by industry affiliated authors *vs.* theoretical articles by college-affiliated authors) utilizing statistical analysis with the chi-square test. The results are discussed and evaluated.

PROCEDURE

To achieve the longest period of coverage, the journal articles were chosen from 1963 journal volumes, so that the citations to them could be looked up in all the yearly cumulations of the "Science Citation Index" (SCI) which began regular publication in 1964.

One of the shortcomings of the citation study by Brookes[1] is that consideration was not made of the increase in publication with each year. This defect is pointed out by Line,[2] who suggested that correction factors be applied to compensate for the increased obsolescence of the particular subject field, as well as the decay of the literature in the field. The desirable consideration of growth factor is stressed, which considers how many articles are published in successive years, but this is very difficult to determine for a particular subject field. A premise is that, if the likelihood of citation is the same for successive years, the earlier years will have fewer citations because there is less literature.

The problem is reflected in the SCI, in that the coverage of the SCI was increased each year; that is, the base of source journals from which citing articles were taken was expanded. In addition, the SCI revealed (in 1971) the results of the study in which the citation frequency varied from year to year, generally increasing by small amounts each year. Although this citation frequency is helpful, it does not consider in its computation the articles that were published that year but which were not cited at all; the figures merely reflect the citations per paper which appear in the SCI. These data are cited in Table I. This phenomenon presents a number of problems. It might seem an easy matter to simply adjust the citation data for each year based on the fraction of total coverage in that year. Although the calculations would be trivial, such manipulations would not really reflect the situation which the statistics suggest. In the present study, the field of chemistry was examined. Chemistry is an old discipline, and the journals for the field have been well established for a long time. Most of the additional coverage of the SCI reflects the birth of new journals which relate to newly evolved fields in science, for example, the whole body of new literature which has been generated by the popular interest in ecology. Further, new disciplines are created by increased specialization and merging of old disciplines, such as (some years ago) when biochemistry came into its own. As mentioned previously, the SCI coverage also extended to some areas of the social sciences, which have little effect on the field of chemistry. A list of the specific journals added was included in the 1971 cumulation: it contained no journals that seemed to be concerned with chemistry.[4] Also, many non-journal sources were added through the years, such as government reports and publications. So the mere yearly figures of "journals covered" is really not an accurate representation of the effect of the increased coverage in the field of chemistry, and it would be misleading and probably counterproductive to apply a corrective factor to the citations collected based on the coverage for that year. One possible procedure to account for the increase in articles would be to list all the journals which contain the *citing* articles in the 1964 cumulation, and then when counting citations in future cumulations only count those that appeared in the

Table I

Year	Journals Covered	Citation Frequency[3]
1964	700	1.60
1965	1147	1.65
1966	1573	1.65
1967	1711	1.66
1968	1968	1.67
1969	2180	1.67
1970	2192	1.73
1971	2287	1.76

1964 edition, thus eliminating new journals in chemistry. One possible problem with this approach is that there may be some chemistry journals that were actually covered in the 1964 edition, but from which no citing articles appeared in the random test sample. Though the SCI generally includes a list of journals covered in the front of each cumulation, it does not include a list of which ones were added each year (except for 1971). Another problem is that citations to chemistry articles do not always appear in chemistry journals.

For these reasons, it was decided not to apply a corrective factor to the figures tabulated. In any case, the comparisons between the behavior of the various categories of articles would not be affected since they would all be subject to the same situation each year of the coverage of the study.

SELECTION PROCEDURES

Journal Selection Procedure. In order to consider significant articles which would be likely to be cited, thereby resulting in an adequate data base and conclusive results, it was necessary to select a certain number of the "top" journals in the field of chemistry. Consultation with reference librarians at the chemistry library of the University of Maryland failed to produce satisfactory results, and, although some journals were considered better than others, reluctance to be specific was evidenced by all those interviewed. Fortunately, the librarians of the Science and Technology Room of McKeldin Library suggested a recent journal article which listed the major articles in various fields of science. The article[5] examined cross-citing among 275 journals in mathematics, physics, chemistry, biochemistry, and biology, data being drawn from the *Journal Citation Index* (a file derived from the "Science Citation Index"). The purpose of the article was to present the hierarchical relationships between the various fields examined in the study, but in the course of this examination a list of the "major journals" for each of the fields was compiled. Since the majority of the study utilized data from 1969, the journals are considered "major journals for 1969," but in view of the relative stability of the field, such a list is an excellent tabulation and indication of the most significant journals from 1964 to 1971.

From the list of ten journals, two were eliminated from consideration for this study. The author's lack of familiarity with German resulted in the elimination of *Chemische Berichte*. To prevent the introduction of counter-productive factors, types of literature other than formal, standard journal articles were not considered. For this reason *Tetrahedron Letters* was eliminated from the study, as was Part B of the *Journal of Polymer Science*, it being thought that the "letters" genre of scientific writing would not conform to the type of article sought. In this vein, Letters to the Editor and Book Reviews were not considered, nor were the special book-length features with many chapters, by one author, which take up an entire journal issue. In addition, the *Journal of Chemical Physics* was evaluated by the study to be a cross-field journal. In order that only articles exclusively about chemistry be included in the study, it too was eliminated from consideration.

After these deletions, the following list of source journals remained:

1. *Acta Crystallographica*
2. *Analytical Chemistry*
3. *Journal of Organic Chemistry*
4. *Journal of Physical Chemistry*
5. *Journal of Polymer Science* (Part A only)
6. *Journal of the American Chemical Society*
7. *Journal of the Chemical Society*

Why these journals are among the "major" journals in the field is a study in itself. Factors considered would be the fact that large journals tend to be cited out of proportion to their size. Correlations have also been found between the citations per paper and total number of papers, but the results of such studies[6] apply to authors rather than journals.

Journal Article Selection Policy/Procedure. To collect the data needed for the study, it was necessary to devise a procedure for determining whether an article is theoretical or practical, and by a college author or by an industry author. For the latter case, all that was necessary was to note the author's affiliation as included under his name in the article. To maintain uniformity of data, articles by authors with dual affiliations were not considered. Also, articles by more than one author, the authors of which had different affiliations (*e.g.*, one in industry, the other at college), were likewise not considered. Institutions of nebulous identity in terms of industry *vs.* college, such as certain research institutes and governmental agencies, were not considered either.

For example, in 1963 the Oak Ridge National Laboratory in Tennessee was operated by the Union Carbide Corporation for the United States Atomic Energy Commission. Sometimes the institution itself was unfamiliar and its identity difficult to determine; in the British *Journal of the Chemical Society* the "Ltd." following the organization helped as a key to classification.

Determination of whether an article is practical or theoretical was more difficult, a skimming of the text of the article usually being necessary to make the decision. Sometimes the affiliation of the journal article writer aided classification, such as one who was with the Department of Applied Chemistry, Nagoya University, Japan.

No articles in languages other than English were considered.

Categories Determination. The object of this study was to compare the behavior of citations to theoretical *vs.* practical articles, college author articles *vs.* industry author articles, and combinations thereof. The procedure was hypothesis generation rather than hypothesis testing. To procure the necessary data, the following categories suggested themselves:

1. Theoretical articles by college-affiliated authors
2. Theoretical articles by industry-affiliated authors
3. Practical articles by college-affiliated authors
4. Practical articles by industry-affiliated authors

It was determined to select 11 articles in each of these four categories for each of the seven source journals. This selection procedure resulted in the following data of source articles:

No. of articles	Type
154	Theoretical
154	Practical
154	College author
154	Industry author
77	Theoretical/college author
77	Theoretical/industry author
77	Practical/college author
77	Practical/industry author
308*	*Total number of articles

Table II

Comparison (category of longer lifetime listed first)	SCA per SA 1	SCA per SA 2	x^2	Level of significance	Av age 1	Av age 2	Av no. of citations 1	Av no. of citations 2
1. Practical *vs.* theoretical	59.2	52.4	9.4546	...	4.48	4.21	13.23	12.44
2. College *vs.* industry	67.8	43.8	83.7021	0.001	4.45	4.20	15.25	10.42
3. Practical/college *vs.* theoretical/college	68.7	67.0	14.4261	...	4.66	4.25	17.08	15.75
4. Practical/industry *vs.* theoretical/industry	49.8	35.1	16.4108	0.050	4.24	3.85	11.73	9.12
5. Practical/college *vs.* practical/industry	68.7	49.8	26.7918	0.001	4.66	4.24	17.08	11.73
6. Theoretical/college *vs.* theoretical/industry	67.0	35.1	77.5329	0.001	4.25	3.85	15.75	9.12
7. Theoretical/college *vs.* practical/industry	67.0	49.8	26.4859	0.001	4.25	4.24	15.75	11.73
8. Practical/college *vs.* theoretical/industry	68.7	35.1	68.3858	0.001	4.66	3.85	17.08	9.12

Once the 308 articles from the 1963 journals were selected, it was then necessary to look up each of the articles in each of the yearly cumulations of the "Science Citation Index" from 1964 to 1971, inclusive, to determine the number of citations per year to each of the source articles. Certain editorial policies were administered: self-citations (citations in which an author cites himself) were not included in the tally; also, citations to an author by another member of the author's research team or his colleagues were avoided where possible, although this was often difficult to determine.

EXAMINATION OF DATA

To facilitate manipulation of the data, A COBOL program was written and executed to perform the necessary functions. Appendices 1 and 2 contain the source-journal data base, appropriate computations, and histograms.

Certain comparisons between the results for each category suggested themselves, and these comparisons were made using the chi-square test. The computations and nature of the comparisons involved are given in Appendix 3.

RESULTS AND ANALYSIS

Examination of the data in the Appendices show that the original hypothesis, namely that theoretical articles have a longer lifetime than practical articles, is supported by the findings of this study. There are a number of possible explanations for this result. The very notion of practical and theoretical might be fallacious. One must ask the question, practical for whom, in making such a determination, a question which defies quantitative evaluation. For example, one theoretical article examined in the study discussed the radiolysis of carbon dioxide, and although the scope of the article seemed to be quite theoretical, some years later a chemist in industry cited the article in a very practical article about distilling liquor. The distinction becomes hazy, because one aspect of a theoretical article is that it should be applied to practice somehow; it should have some useful purpose. Furthermore, what is theoretical today might have a long-range practical application; this might or might not be realized by the author, and might be implicit in what otherwise seems to all but the most knowledgeable reader as a theoretical article. This ambiguity can be seen in the results of the chi-square test comparison between practical and theoretical articles, which did not support the hypothesis that the populations are different. Also, there are problems unique to the discipline of chemistry which contribute to the confusion. Since there is not the preponderance of general laws in chemistry as there are in other disciplines, a theoretical article about the structure of a molecule, for example, becomes the practical guide to which chemists refer and by which they formulate their activities in the laboratory. Once such a molecular structure has been thus examined, it is generally held to be true indefinitely (unless of course the experimental procedure is found to be faulty for some reason).

At this point it would be useful to refer to Table II for discussion of the results of the comparisons made. The Sum of Citation Ages per Source Article (SCA per SA) for each category gives weight to the sheer number of citations, whereas the Average Age computations reveal the age distribution over the period in an attempt to compare the shapes of the respective curves generated by the yearly citation tables. Generally, the predominant influence seems to be whether the article is college or industry, not whether it is judged to be practical or theoretical. These characteristics tend to most greatly influence the combinations of the resulting comparisons. Furthermore, a hypothesis can be generated about the impact of theoretical articles by industry-affiliated authors on the chemistry field: namely, that the impact is significantly lower than other types of articles. Such articles have a significantly lower lifetime, judging from the computations made, and one can only hypothesize that such articles are not effective in stimulating interest in theoretical matters. Perhaps chemists in industry regard such articles as of less value than practical articles, or chemists at college regard theoretical articles by college chemists more highly than those by industry chemists, both hypotheses resulting in the articles being more neglected. Another possibility is that the articles are just bad articles.

The greatest span of the compound comparisons was noted between practical/college (the category with the highest lifetime) and theoretical industry (the lowest). This result owes to the fact that college lives longest, overall, and industry lives shortest. Although the hypothesis that the populations of practical and theoretical are different was not supported by the chi-square test, whether an article is practical or theoretical can be an influencing factor in the compound comparisons. Even in view of the results, however, the significance of whether an author is college or industry can be questioned, since many chemists migrate; one year a chemist might be affiliated with M.I.T., whereas the next year he is working for DuPont. Thus the validity of the comparison could even be questioned; still, the author could continue to write theoretical (or practical) articles regardless of his affiliation, but the climate in industry might be more conductive to practical articles and the climate at college might be more conductive to theoretical articles.

The results of the initial comparisons are reflected in the compound comparisons. That is, since practical has a longer life than theoretical overall, this pattern repeats within industry and within college. The same is true of college *vs.* industry. The degree of difference varies, but generally the range of significance is similar and further hypotheses would require more data and different kinds of comparison.

LITERATURE CITED

(1) Brookes, B. C., "Obsolescence of Special Library Periodicals: Sampling Errors and Utility Contours," *J. Amer. Soc. Inform. Sci.*, **21**, 320–329 (1970).

(2) Line, M. B., "The 'Half-Life' of Periodical Literature, Apparent and Real Obsolescence," *J. Doc.*, **26** 46–54 (1970).

(3) "Science Citation Index 1971," Vol. V., Institute for Scientific Information, Philadelphia, Pa., 1972, p 3.

(4) *Ibid.*, p 44.
(5) Narin, F., *et al.*, "Interrelationships of Scientific Journals," *J. Amer. Soc. Inform. Sci.*, **23,** 323-331 (1972).
(6) Cole, S., and Cole, J. R., "Scientific Output and Recognition: A Study of the Operation of the Reward System in Science," *Am. Sociological Rev.*, **62,** 377-390 (1967).

RELATED REFERENCES

Brown, G. W., Miller, J. G., and Keenan, T. A., "EDUNET: Report of the Summer Study on Information Networks," Conducted by Interuniversity Communications Council, Wiley, New York, N. Y., 1967.
Burton, R. E., and Kebler, R. W., "The 'Half-Life' of Some Scientific and Technical Literature," *Amer. Doc.*, **11,** 18-22 (1960).
Carter, L. F., *et al.*, "National Document-Handling Systems for Science and Technology," Wiley, New York, N. Y., 1967.
De Solla Price, D. J., in "Communication in Science: The Ends—Philosophy and Forecast," A. De Reuck and J. Knight, Ed., J. and A. Churchill, London, 1967.
De Solla Price, D. J., "Networks of Scientific Papers," *Science*, **149,** 510-515 (1965).
Garfield, E., "Citation Indexes for Science," *Science*, **122,** 108 (1965).
Garfield, E., and I. H. Sher, "New Factors in the Evaluation of Scientific Literature through Citation Indexing," *Amer. Doc.*, **14,** 195-201 (1963).
Hays, W. L., "Statistics," Holt, Rinehart and Winston, N. Y., 1963.
Hodges, J. L., and Lehmann, E. L., "On Basic Concepts of Probability and Statistics," 2nd ed, Holden-Day, San Francisco, Calif., 1970.
Keen, E. M., "Citation Indexes," *ASLIB Proc.*, **16,** 246-51 (1964).
Lancaster, H. O., "The Chi-squared Distribution," Wiley, New York, N. Y., 1969.
Loosejes, T. P., "On Documentation of Scientific Literature," Translated by A. J. Dickinson, CT. Archon Books, London and Hamden, 1967.
Smith, J. R., and East, H., "Information Services in Physics," Communication in Science: Documentation and Automation, A. De Reuck and J. Knight, Ed., J. and A. Churchill, London, 1967.
Velke, L., "The Use of Citation Patterns in the Identification of Research Front Authors and Classic Papers," *ASIS Proc.*, **7,** 49-51, (1970).

APPENDICES (Available on request)

1. Data base: list of 308 journal articles cited
2. Cumulation/computation of data; histograms
   2.1 Citations, cumulations tables/histograms
   2.2 Sum of Citation Ages per Source Article, computations/histograms
   2.3 Average Age/percentage of citation computations/histograms
3. Chi-square test computations

Revised Nomenclature for Highly Fluorinated Organic Compounds

JOHN A. YOUNG

Department of Chemistry, Universidad Autonoma de Guadalajara, Guadalajara, Jalisco, Mexico

Received February 4, 1974

The system described herein has been adopted by the American Chemical Society as authorized nomenclature for highly fluorinated organic compounds. Its main features are the introduction of the symbol "*F*" to convey the sense of "perfluoro" and the prefix "hydryl" to describe solitary hydrogen atoms in an almost completely fluorinated structure, with codification of standard usage for these prefixes.

There has been no change in authorized nomenclature for organic fluorine compounds since 1952, when the conventions "*H*" and "perhalo" were adopted.[1] Prior to that time, it had been necessary to name and locate each fluorine atom in a molecule such as $H(CF_2)_5CF_3$. The "*H*" convention simplified and shortened nomenclature in this and similar cases but unfortunately could not be used together with "perfluoro," so that, for example, $H(CF_2)_6Cl$ remained 6*H*-1,1,2,2,3,3,4,4,5,5,6,6-dodecafluorochlorohexane. Moreover, the term "perfluoro" was employed erroneously and ambiguously in spite of its wide usage, since it sometimes included a functional group and at other times did not; for example, both $(CF_3)_2NH$ and $(CF_3)_2NF$ were referred to in the literature as perfluorodimethylamine. The new system replaces "perfluoro" with "*F*", which conveys the same sense of complete fluorination without the use of locants, and replaces "*H*" with "hydryl," a collective prefix which can be applied in conjunction with "*F*" to functional compounds.

Established systems of organic chemical nomenclature are all based on the prevalence of hydrogen. The total number of hydrogen atoms in a molecule is not stated but must be found by satisfying all the conditions stipulated in the name concerning atomic identity, valence state, and structure. Any atom whose identity is neither stated nor implied in any part of the name is assumed to be a hydrogen atom. The main principle of the rules presented here is the transfer of this assumed prevalence from hydrogen to fluorine, by use of the symbol "*F*". When this symbol is used, any atom attached to carbon whose identity is neither specifically stated nor implied in any part of the name is assumed to be a fluorine atom. As the parent "*F*" compound or radical, by this definition, contains no hydrogen atoms attached to carbon atoms in the named backbone, any such hydrogen atoms actually present cannot be regarded as residual but must be named as substituents, and the total number of fluorine atoms is not stated but results from satisfying all the stipulated conditions of atomic identity, valence state, and structure. In other words, substituent atoms or groups are regarded as having been substituted for fluorine, not for hydrogen.

These rules are concerned with what is tacit within a name, that is, the nature of unspecified atoms, rather than with structural features. Consequently, once the implication of the symbol "*F*" is made, all other features of the molecule can be described, in general, by conventional naming practices.

The rules deal with three areas: (a) definition of the symbol "*F*" and delineation of its proper usage, (b) treatment of fluorocarbon derivatives containing heteroatoms, and (c) designation of hydrogen atoms where only a few of these are present. The present rules deal only with compounds