The Law of Exponential Growth: Evidence, Implications and Forecasts

JEAN TAGUE
JAMSHID BEHESHTI
LORNA REES-POTTER

THE NOTION THAT KNOWLEDGE grows exponentially seems to have first appeared in a short story by Sir Arthur Conan Doyle, "The Great Keinplatz Experiment," which contains the statement, "Knowledge begets knowledge as money bears interest."[1] Thus, knowledge growth is likened to compound interest—the increase at any time is a fixed percentage of the current amount. This type of growth is described mathematically by an exponential function. If F(t) represents the size at time *t*, the exponential function, or law, may be expressed as

$$F(t) = ae^{bt} \tag{1}$$

where *a* is the initial size—i.e., at time t =0—and *b*, the continuous growth rate, is related to the percentage by which the size increases each year (or other appropriate time unit). Specifically, this percentage is given by

$$r = 100(e^{b}-1), \text{ or, approximately, } r = 100b.$$

For example, if the amount of knowledge at some initial time is a =10,000 and the growth rate is approximately r =10 percent, then after 10 years the amount of knowlege will be

$$F(10) = 10{,}000e^{0.1(10)} = 27{,}183.$$

After 100 years the amount will be

$$F(100) = 10{,}000e^{0.1(100)} = 220{,}264{,}660.$$

Jean Tague is Professor, School of Library and Information Science, and Jamshid Beheshti and Lorna Rees-Potter are doctoral students, University of Western Ontario, London.

Another quantity that is of interest with respect to exponential growth is doubling time: the fixed period of time in which the size of the literature doubles. Doubling time is given by

$$d = \log_e 2/b.$$

For the above example, the amount of knowledge doubles every $d = 0.693/0.1 = 6.93$ years.

Not all writers agree on the exponential nature of this growth. Popper says "the growth of knowledge...is not a repetitive or cumulative process, but one of error elimination."[2] Similarly, Rescher comments: "Science progresses not additively but largely subtractively. Today's major discoveries represent an overthrow of yesterday's."[3] Price[4] has brought the idea of exponential knowledge growth in the sciences to the attention of a wide audience. He looks at various indicators of growth, including the number of scientists, number of scientific journals, number of scientific abstracts, and amount of scientific expenditure. For the scientific literature, he found a growth rate of approximately 5 percent over the past two centuries, corresponding to a doubling time of fifteen years. Growth of knowledge must be distinguished from growth of the literature or growth in number of publications. The former is a more abstract concept and hence not so directly assessed. In bibliometrics, growth in number of publications is sometimes taken as a measure or operational definition of growth of knowledge. There are, however, other points of view. Rescher defines the λ-quality level, $0 < \lambda \leq 1$, of a publication or finding as follows: if there are $F(t)$ publications in all at time *t*, then there will be $[F(t)]^{\lambda}$ publications at the λ-level. He characterizes specific values as follows:

$\lambda = 1$ at least routine
$\lambda = 3/4$ at least significant
$\lambda = 1/2$ at least important
$\lambda = 1/4$ at least very important
$\lambda = 0$ first-rate

For first rate contributions ($\lambda = 0$), the number of publications is $\log F(t)$. Rescher points out that the value of ½ corresponds to Rousseau's law, which states that the number of important contributions is the square root of the total number of contributions. Thus, if the size of the literature is 1 million publications, in terms of Rescher's λ-levels, there would be:

1,000,000 at least routine publications
31,623 at least significant publications
1,000 at least important publications

32 at least very important publications
14 first-rate publications

If the total literature (assuming anything published is at least routine) is growing exponentially with a doubling time *d*, then the literature of λ-quality, for $\lambda > 0$, is growing exponentially with the doubling time of d/λ. Thus, as one ascends the quality scale, exponential growth slows down. For first-rate literature, exponential growth breaks down completely and there is merely a constant increment in each time period. In this case the growth function is linear, i.e., the number of first-rate publications at time *t* is given by

$$F_0(t) = \log a + bt$$

when the total number of publications is given by (1). Here, *b* would represent the constant increment. In the earlier example, in which the doubling time was 6.93 years, the corresponding doubling times for each λ-level group of publications would be

9.24 years for at least significant publications,
12.60 years for at least important publications,
27.73 years for very important publications.

The number of first-rate publications at time *t* would be given by the function

$$F_0(t) = 9.21 + 0.1t$$

That is, there is only one additional first-rate publication every ten years.

Exponential increase occurs when there are no limits to growth. However, if there is some limitation, intellectual, physical, or economic, on the size of the literature, then other functions, such as the logistic, may be more appropriate. Price points out that organisms in a closed environment (e.g., fruit flies in a bottle) tend to follow a logistic rather than an exponential growth function. The logistic curve is characterized by a lower limit (usually 0) and an upper limit or ceiling, beyond which size cannot grow. The equation for the logistic curve is

$$F(t) = \frac{k}{1 + ae^{-bt}}$$

where F(t) represents the size at time *t*, and *k* the ceiling. The shapes of the logistic curve and exponential and linear ones in the same range are shown in figure 1. The curve is symmetrical about the point of inflection at

$$t = \frac{\log a}{b} = t'.$$

If $t < t'$, the growth rate is increasing; if $t > t'$, the growth rate is decreasing. Using the previous hypothetical example, if size at the initial time t=0 is 10,000 publications, the initial yearly growth rate is 10 percent and the upper limit is 300 million publications, then the appropriate logistic function is

$$F(t) = \frac{300{,}000{,}000}{1 + 29{,}999e^{-0.1t}}.$$

After ten years the size of the literature would be 27,181 publications, i.e., almost the same as under exponential growth. However, after 100 years, the size would be only 127,013,560, instead of the 220,264,660 publications which would be obtained with exponential growth.

The growth pattern of subfields of knowledge or research areas may be different from that of the parent field. Crane[5] suggests that some subfields show "the first three stages" of a logistic pattern. These fields are diffusion of agricultural innovations, 1941-66 (sociology); and theory of finite groups, 1934-68 (mathematics). Her characterization of logistic growth is not strictly accurate. It involves four stages: a slow start, a period of exponential growth, a period of linear growth, and then a period of slow, irregular growth. However, as indicated above, the logistic curve is perfectly symmetrical on either side of the midpoint with the growth rate always increasing before the midpoint and always decreasing after the midpoint, but never constant or linear. In fact, the growth curves shown for Crane's two subfields could equally well be described as exponential followed by linear. This pattern was also found by Lawson and others[6] in the energy analysis subfield. The closest approximation to a true logistic curve seems to be the growth curve of the coal gasification literature for the period 1965-75, as described by Frame, et al.[7]

In two other fields, invariant theory (1887-1941) and reading research (1881-1957), Crane found a linear growth pattern. Sullivan found a similar pattern in the physics literature, both experimental and theoretical, concerned with weak interactions for the period 1950-72. Menard found linear growth in the subfield of optics, but in three other subfields of physics he found exponential growth, though at differing rates: nuclear physics has doubled every four or five years since 1920 and solid state physics since 1950; acoustics, on the other hand, had a doubling time of forty years prior to World War II, but since then has been doubling at normal rates—i.e., every fifteen years.[8]

Menard distinguishes three types of subfields: stable fields, which tend to grow linearly or exponentially at very slow rates; growth fields, which grow exponentially at fast rates; and cyclic fields, which fluctu-

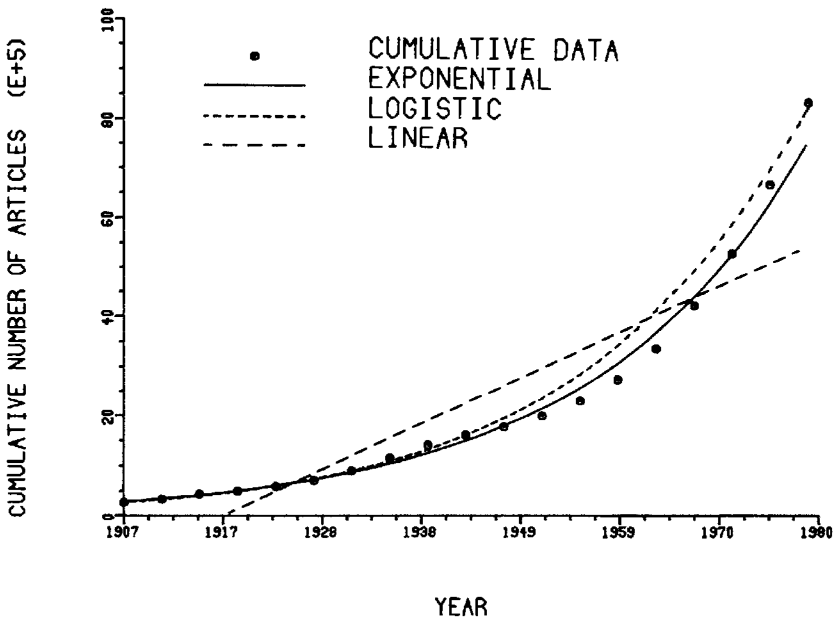

Fig. 1. Cumulative numbers of *Chemical Abstracts* fitted by least-squares to linear, exponential and logistic functions.

ate, with stable and growth periods alternating. An example of a stable field would be vertebrate paleontology, described by Menard. An example of a growth field would be activation analysis (chemistry), described by Braun,[9] for which doubling time over the period 1935-75 has been three years. An example of a cyclic field—liquid crystals—was presented by Bottle and Rees.[10] During the period 1888-1974, the number of publications increased to a peak in 1910, then decreased and lay dormant in the 1930s and 1940s, then increased exponentially in the 1960s. Menard suggests that the overall growth rate of a discipline varies at different times depending on the proportion of papers from stable, growth and cyclic fields.

Goffman's epidemic model is, to some extent, similar to Menard's cyclic model. Scientists are classified as: (1) infectives—those currently publishing in the field, (2) removals—those who have published in the past, and (3) susceptibles—those who may publish in the future. If S(t), I(t) and R(t) represent, respectively, the number of susceptibles, infectives, and removals at a point in time *t*, then the change in these functions can be described by a set of differential equations and a threshold level determined for the number of susceptibles required to produce an epidemic. The constants in these equations represent the rate of infection, the rates at which susceptibles and infectives are removed, and the rates at which new supplies of infectives and susceptibles enter the population. The model has been applied to the research literature of mast cells;[11] shistosomiasis, 1862-1962;[12] symbolic logic, 1847-1962;[13] and polywater, 1962-74.[14] The curves for the first two literatures display the usual exponential pattern; symbolic logic literature is cyclic, with peaks in 1907, 1932 and 1957; and polywater literature has a single peak in 1970.

The epidemic model is difficult to evaluate because of the indefiniteness in its presentation and applications. In no case are all three functions S(t), I(t) and R(t) stated explicitly as functions of time, although an exponential form is suggested for I(t). Also, the constants required in the differential equations are not all estimated from the empirical data. The impression is that any kind of cyclic or exponential growth pattern is compatible with the epidemic model.

One general problem in describing the literature growth of a subfield is that it is difficult to determine when the subfield first arose from its originating field. As Menard has pointed out, indexes and abstract journals do not ordinarily create new classes or subheadings until after the first 100 or so papers have appeared. Eventually, if the subfield becomes very large, it will split into two or more subfields. Increasing specialization is the response of scientists to an increasing literature burden. However, recent investigations by Small indicate it may be possible to identify specialties by means of cocitation-based content analysis.[15]

The Evidence

What is the evidence for exponential growth? The answer depends on what one is counting and when.

Knowledge growth may mean literature growth—increase in the number of publications in a field—or information growth—increase in the number of ideas in the field. As Gilbert[16] has pointed out in connection with indicators of scientific growth, the use of the former as a measure of the latter assumes, first, that all knowledge is contained in the published literature, and second, that every paper contains an equal amount of knowledge.

Even if number of publications (where the word *publication* is used in a broad sense to mean anything in the form of text) is a reasonably valid approximation of the amount of knowledge, the reliability of counts of publications in specific fields must be questioned. Usually, these are based on items in the standard abstracting journal for the field. Moravcsik[17] has pointed out that many scientific communications do not appear as articles in scientific journals, the primary source of materials for the abstract journals. Abstract journals are biased geographically and linguistically; they do not include material in near-print form, material which results from military or proprietary research and is not published in the open literature, or informal person-to-person communication. Although the ideas in these other materials may appear eventually in print, it is difficult to assess the number that do not.

Bearing in mind the limitations of these data, let us, however, examine the growth of the literature as revealed by counts of the number of abstracts in some of the major abstracting journals. The chemical literature has been analyzed more than any other, probably because of the wide coverage of *Chemical Abstracts* and the stability of its growth pattern. Figure 1 shows the cumulated number of chemical abstracts up to 1979, together with the best-fitting linear, exponential and logistic curves. By a cumulated curve is meant one in which the number of abstracts is cumulated or summed from year to year, beginning at a specified point in time—in this case, 1907. Best fit is defined by the least-squares criterion. In looking at the literature of literature growth, one is struck by the absence of data fitting by least squares. Most exponential growth rates seem to be determined by eye from the empirical plots. Usually, the reader can determine empirical values only approximately from the plots rather than exactly from a table. It is thus difficult to check on the specified growth rates, doubling times and other characteristics deduced by the author. The counts upon which the figures in this paper are based are given in the appendix.

May[18] has pointed out that by beginning a cumulated curve in a specific year such as 1907, the earlier literature is ignored. This usually results in an overestimation of growth rates. For example, if the cumulated totals for the mathematics literature are begun in 1920 rather than in 1868, the growth rate increases from 2.5 percent to 4.6 percent. May's method for including the earlier literature is to fit the noncumulated annual counts of publications to an exponential curve. This curve is then integrated to obtain the corresponding cumulated curve. The continuous growth rate (b in equation 1) will be the same for both curves, but the constant factor (*a* in equation 1) will change. For example, applying May's method to the annual noncumulated output for *Chemical Abstracts* 1907-79, one obtains the exponential curve:

$$f(t) = 12{,}061\ e^{0.046(t-1906)}$$

If this function is integrated from $-\infty$ to 1907, the estimated cumulated number of chemical publications prior to 1907, i.e., 262,196, is obtained. This number is then added to the cumulated number of publications since that time, as determined from *Chemical Abstracts* counts, to obtain the data points in figure 1. The three theoretical curves are the least-squares exponential, linear and logistic fits to these points. The corresponding functions and multiple squared correlation coefficients are given in table 1. The squared correlation coefficient represents the proportion of the variation of cumulated size values which can be explained by the theoretical function. The algorithm developed by Oliver[19] was used in an attempt to find a least-squares fit to the logistic curve, but unfortunately did not converge. The function given is thus only an approximation to the least-squares solution.

TABLE 1

FUNCTIONS APPROXIMATING THE CUMULATIVE NUMBER OF CHEMICAL ABSTRACTS, 1907-79

Type	*Function*	R^2
Linear	$F(t) = -999{,}000 + 88{,}013(t-1906)$	0.811
Exponential	$F(t) = 282{,}546.94e^{.044842(t-1906)}$	0.995
Logistic	$F(t) = \dfrac{44{,}751{,}400}{1 + 170.743e^{-.0499885(t-1906)}}$	0.986

For the *Chemical Abstracts* data, 1907-79, the exponential growth rate is thus 4.5 percent, corresponding to a doubling time of fifteen years. For the linear fit, the constant increment is 88, 013 papers per year. The midpoint of the logistic fit is at the year 2008, and the upper limit for this function is 44,751,400 papers.

To compare the growth of the chemical literature with that in other fields, annual counts of the number of abstracts from 1960 to 1979 were recorded for the following journals: *Science Abstracts* (physics, electrical engineering, computers, and control), *Biological Abstracts, Chemical Abstracts, Psychological Abstracts, Library and Information Science*

Abstracts, International Political Science Abstracts, Historical Abstracts, and *Sociological Abstracts.* Figure 2 shows cumulated number of abstracts in *Chemical Abstracts, Science Abstracts* and *Biological Abstracts*, 1960-79; figure 3 showns the same data for *Sociological Abstracts, International Political Science Abstracts* and *Historical Abstracts*; figure 4, the same data for *Psychological Abstracts*; and figure 5, the same data for *Library and Information Science Abstracts*. Groupings were determined, in part, by the scale of the vertical axis, and in part by similarities in subject matter. In these cases, no correction was made for pre-1960 literature, so that the data points shown in figures 2, 3, 4, and 5 show cumulations relative to 1960 only. By fitting exponential functions to both the noncumulated and cumulated values, using May's method described earlier, it was possible to obtain growth rates either incorporating or ignoring the pre-1960 literature. Fits were also made just to the 1970-79 figure to determine if growth was changing in the seventies.

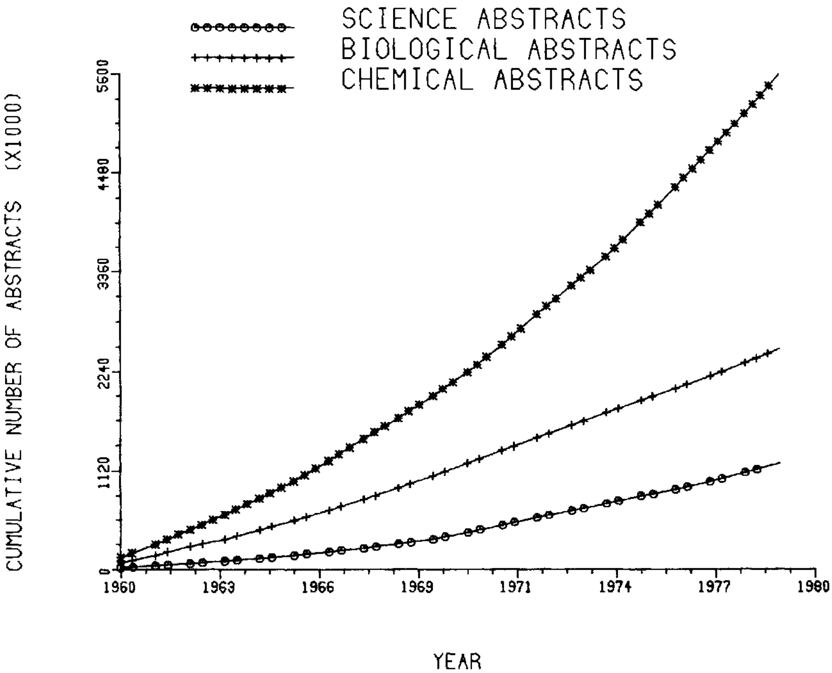

Fig. 2. Cumulative numbers of abstracts in three abstract journals, 1960-79.

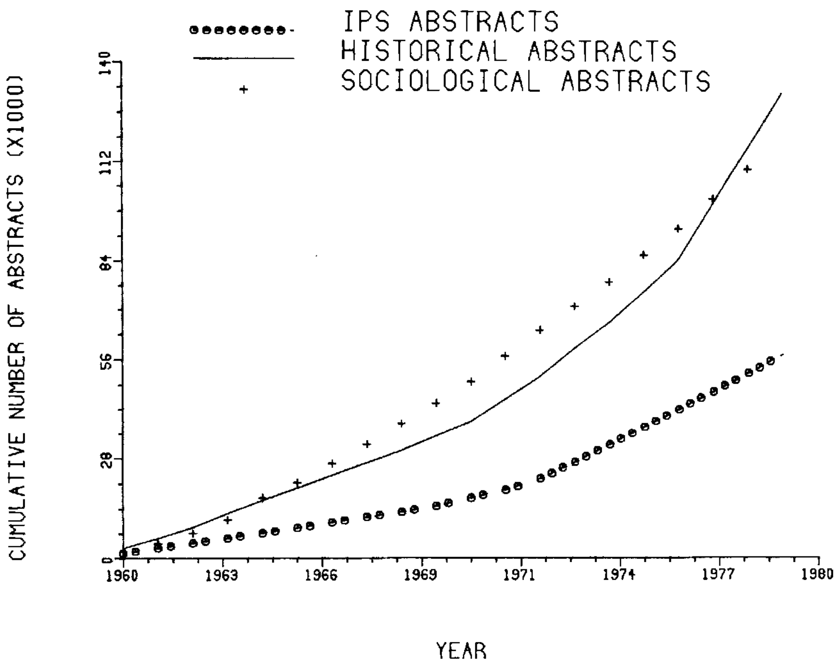

Fig. 3. Cumulative numbers of abstracts in three abstract journals, 1960-79.

The annual growth rates for the two periods, 1960-79 and 1970-79, based on cumulated and noncumulated figures, are shown in table 2. An examination of these indicates that in the seventies, for the most part, growth is slowing down. Rates are generally higher in the social sciences than in the physical and biological sciences, but it is not clear whether this difference is due to an increase in the social science literature or a change in coverage of the abstracting journals. As far as chemistry is concerned, Baker, in a review of *Chemical Abstracts* growth rates,[20] says that the journal coverage policy for *Chemical Abstracts* has not changed in twenty-five years, although that for patents has changed. The smaller growth rates obtained when the noncumulated values are taken into account are consistent with May's predictions. Only in one out of sixteen cases, *Historical Abstracts* for 1970-79, are the noncumulated rates greater than the cumulated ones. This anomaly may be due to the strange behavior of *Historical Abstracts* annual production, which increased approximately 60 percent in 1977. Also remarkable is the wide

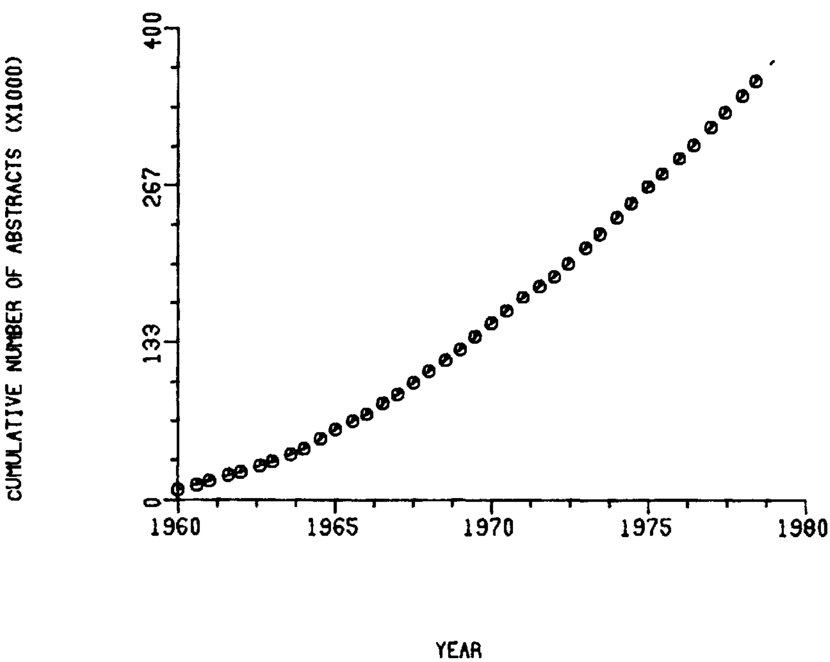

Fig. 4. Cumulative numbers of abstracts in *Psychological Abstracts*, 1960-79.

variation in growth rates from decade to decade and science to science, making questionable such blanket statements as "the scientific literature is growing at 5 percent per year." Also, it is not always clear, when authors are discussing the growth of science, whether just the physical and biological sciences are intended, or the social sciences as well.

The annual and cumulated data for each abstracting journal and for the two time periods were fit to both exponential and linear functions using least-squares procedures. The resulting squared correlation values are given in tables 3 and 4. In all cases, reasonable fits can be obtained to either an exponential or linear function. In all cases except *Library and Information Science Abstracts, International Political Science Abstracts,* and *Historical Abstracts,* the linear fits were better for the 1960-79 data, both cumulated and noncumulated. Thus, growth does seem to be slowing down and moving toward a linear rather than an exponential stage.

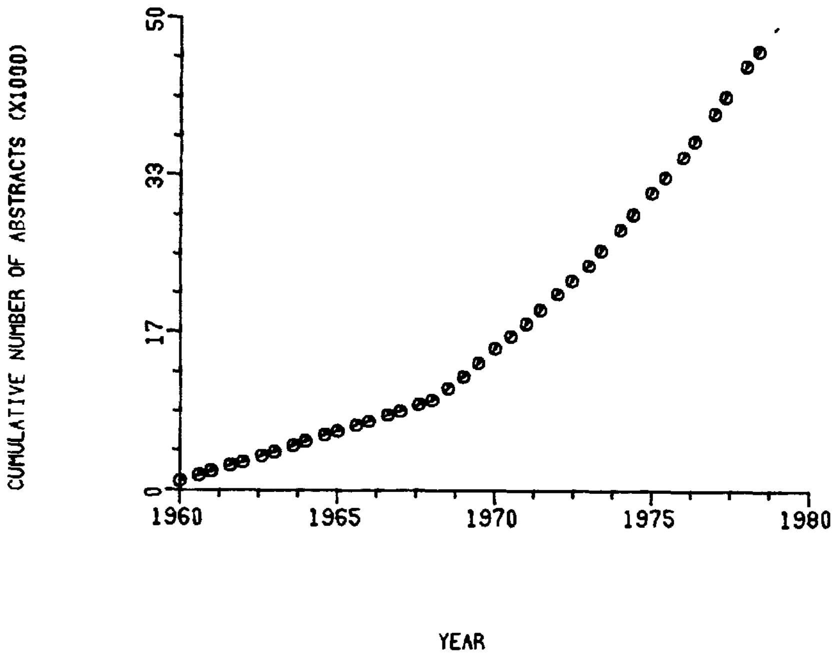

Fig. 5. Cumulative numbers of abstracts in *Library and Information Science Abstracts*, 1960-79.

TABLE 2
ANNUAL GROWTH RATE PERCENTAGES FOR ABSTRACTS IN EIGHT ABSTRACTING JOURNALS, 1960-79

	Annual Growth Rates			
Abstract Journal	*1960-79 Non-cumulated*	*1960-79 Cumulated*	*1970-79 Non-cumulated*	*1970-79 Cumulated*
Science Abstracts	9.0	19.0	2.0	11.4
Biological Abstracts	3.3	15.4	1.0	8.0
Chemical Abstracts	6.2	16.6	4.8	10.1
Psychological Abstracts	7.3	17.8	3.5	10.1
Library and Information Science Abstracts	10.2	18.3	6.4	13.2
International Political Science Abstracts	8.8	16.6	9.8	13.9
Historical Abstracts	9.3	16.7	14.4	13.4
Sociological Abstracts	6.7	19.0	3.3	9.7

TABLE 3

SQUARED MULTIPLE CORRELATION COEFFICENTS FOR LINEAR AND EXPONENTIAL FITS TO CUMULATED NUMBERS OF ABSTRACTS, 1960-79

Abstract Journal	*Linear Fit*	*Exponential Fit*
Science Abstracts	0.959	0.937
Biological Abstracts	0.995	0.883
Chemical Abstracts	0.977	0.911
Psychological Abstracts	0.977	0.925
Library and Information Science Abstracts	0.930	0.960
International Political Science Abstracts	0.923	0.954
Historical Abstracts	0.919	0.940
Sociological Abstracts	0.987	0.879

TABLE 4

SQUARED MULTIPLE CORRELATION COEFFICIENTS FOR LINEAR AND EXPONENTIAL FITS TO NONCUMULATED NUMBERS OF ABSTRACTS, 1970-79

Abstract Journal	*Linear Fit*	*Exponential Fit*
Science Abstracts	0.913	0.910
Biological Abstracts	0.833	0.770
Chemical Abstracts	0.984	0.982
Psychological Abstracts	0.922	0.864
Library and Information Science Abstracts	0.901	0.898
International Political Science Abstracts	0.821	0.853
Historical Abstracts	0.759	0.880
Sociological Abstracts	0.884	0.784

Abstract journal counts are useful for estimating growth within a discipline. However, they cannot be added together to determine overall literature growth because of journal overlap. Some attempts have been made to estimate the total number of journals, but these seem to have a rather low reliability, being heavily dependent on the source of the counts. *Ulrich's International Periodical Directory*, 1979-80, estimated its total coverage to be 62,000 periodicals. Carpenter and Narin[21] used a magnetic tape of all serial publications received by the British Lending Library Division in 1973 and came up with 16,346 journals in the fields of clinical medicine, biomedicine, biology, chemistry, physics, earth and space science, psychology, mathematics, and engineering. An earlier count by Hulme in 1921, based on journals referred to in the *International Catalog of Scientific Literature*, 1908-12, produced 7610 journals (excluding psychology and engineering).[22] Thus, for scientific journals, the recent doubling time appears to be 57 years. A different figure for total number of scientific and technical journals is given by Gottschalk and Desmond of the Library of Congress in 1963.[23] Their figure is 35,000 ± 10 percent, and is based on a perusal of the most comprehensive and recent serial directory for each country. In 1962, Bourne estimated the total number of journals, based on an inventory being performed at the Science and Technology Division at the Library of Congress, as 30,000 to 35,000.[24] The percentage of the literature covered by abstracting journals varies from field to field. Overall, it is about 75 percent, but ranges from 98 percent for chemistry to 50 percent for biology. These percentages were estimated by editors and others knowledgeable in the subject field. Thus, if Bourne's figures are correct, the totals shown in figures 2-5 have varying reliability as measures of the total literature production in a field.

Knowledge, particularly in the humanities, may be better represented by book rather than journal article production. Figure 6 shows cumulated figures for numbers of first-edition titles produced by the principal English-speaking countries, with the exception of Australia, as compiled in the Unesco *Statistical Yearbook*. The data are available for ten consecutive years from 1967 to 1976 for Canada, New Zealand, United Kingdom, and the United States. The data constitute 24.8 percent of the world production of first editions for 1976. Of this figure, 17.2 percent is from the United States, 5.8 percent from the United Kingdom, 1.4 percent from Canada, and 0.4 percent from New Zealand. Unfortunately, Australian figures were incomplete and had to be omitted. Some inconsistencies exist among the various countries. Whereas Canada does not include its government publications in book production figures, 20 percent of the 1976 U.S. data consist of federal government publications. In figure 6, the data will be seen to be linear ($r^2 = 0.998$ *v.* $r^2 = 0.919$ for the exponential function).

Interpretation

To what extent does number of publications actually measure knowledge? Does each publication make a significant and equal contribution to the stock of ideas? One of the few empirical investigations of

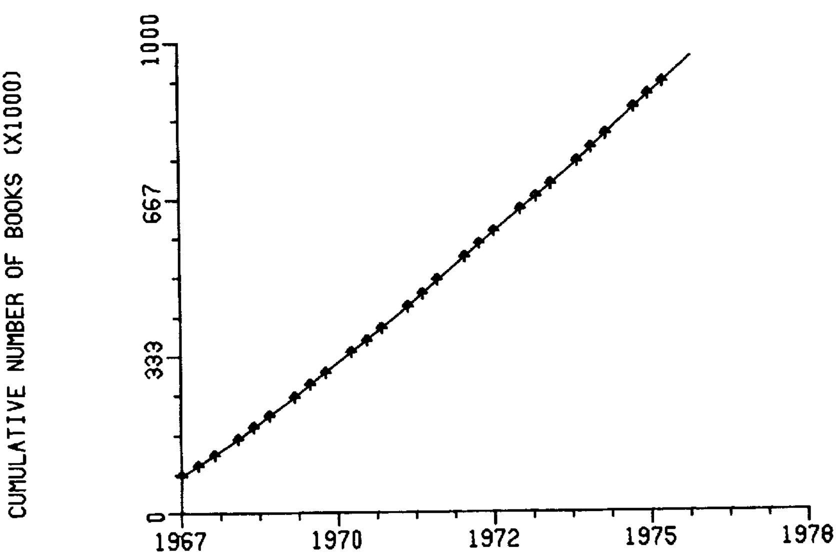

Fig. 6. Cumulative numbers of first editions published in the United States, United Kingdom, Canada, and New Zealand, 1967-76.

this question was carried out by May,[25] who classified mathematical papers on the subject of determinants, as contained in a 1923 bibliography, into six categories: new ideas and results, applications, systematization and history, texts and education, duplications, and trivia. The numbers of articles in each category and percentage of total is shown in table 5. If these numbers are compared with Rescher's λ-quality index and Rousseau's law, it is apparent that, in subject area of determinants at least, there are more than $\sqrt{1995} \approx 45$ important papers and log $(1995) \approx 8$ first-rate papers. However, the discrepancy may arise from the fact that May considers as "literature" only scientific contributions abstracted in professional mathematical journals, but not popularizations and elementary textbooks. Thus, the total number of publications is probably greater than 1995.

May also analyzes individual time trends in each category. New results and ideas are stable, averaging about three per year. Applications

TABLE 5
MAY'S CATEGORIZATION OF THE LITERATURE OF DETERMINANTS TO 1920

Category	*Number of Papers*	*Percentage*
New ideas and results	235	12
Applications	208	10
Systematization and history	199	10
Texts and education	266	13
Duplications	350	18
Trivia	737	37

are closely correlated with new results, with some time lag. Pronounced peaks are observed in texts, publications and trivia. May describes the pattern as follows: "First the basic theory is worked out in close relation to applications. Its successes lead to many textbooks and then to a rush into the field of workers who inevitably lower over-all quality."[26]

Surprisingly, considering its importance to bibliometric approaches to the growth of knowledge, May's study has not been duplicated in other subfields. Of course, such analyses are very time-consuming and require expert knowledge. A criticism can be made that the assignment to categories is very subjective. Also, such a categorization fails to recognize that some duplication is necessary to ensure that new results reach a variety of audiences. However, in general, such analyses can be very revealing.

To investigate the viability of May's approach in another subfield and to familiarize ourselves with its problems, we applied a similar analysis to studies of obsolescence of library materials. The corpus of papers was obtained by checking the heading "Obsolescence of books, periodicals, etc." in *Library Literature* from its first appearance in 1970 and then extending the set to include appropriate references contained in the initial articles. The survey was restricted to English-language items.

Because of the small number of papers, forty-six in all, they were divided into four (rather than six) categories: (1) new ideas and results; (2) new applications; (3) reviews and historical surveys; and (4) popularizations, duplications, trivia. Initially, each paper was categorized by two of the writers independently. Disagreements were then resolved by discussion and more precise definition of the categories. The publication dates ranged from 1944 to 1980. The numbers and percentages for each category are given in table 6. Although not nearly so comprehensive as May's study, these figures do seem to substantiate his finding that new ideas and results (innovations) account for a relatively small percentage (in this case, 28.2 percent) of the total. The variation over time is shown in figure 7. The number of innovative articles remains relatively constant, whereas the total number increases, possibly exponentially, over the time period.

TABLE 6
LITERATURE OF OBSOLESCENCE, 1944-80

Category	*Number of Papers*	*Percentage*	*Number of Authors*
New ideas and results	13	28	11
Applications	11	24	11
Surveys and reviews	3	7	3
Other	19	41	16

It has been suggested by Price and other bibliometricians that the degree to which articles represent innovations can be determined from citation counts. To assess this claim, the number of citations to each of the obsolescence papers published in the period 1944-77 was determined from *Social Sciences Citation Index*. Later papers were not included, as they had probably not yet really entered the citation cycle. Table 7 shows, for each category, the number of papers, the average number of citations per paper, and the minimum and maximum numbers of citations. It is interesting that in category 1, the earliest paper located (that by Gosnell in 1944[27]) received only two citations. Apparently it was ahead of its time. Overall, one must conclude from this brief survey that although citations do give some indication of quality, they can be so used only in an approximate or average way and not for individual papers.

Some historians and sociologists have made similar points about the use of publications as growth indicators and of citations as quality indicators. Moravcsik notes that differences in publication patterns in different countries and different fields make the use of a paper as a unit of knowledge somewhat suspect.[28] Computers may eventually so change the nature of papers and citations that it will no longer be possible to count them in any meaningful way. Also, once a discovery has entered the public domain, e.g., Einstein's equation $E = mc^2$, the original paper is not usually cited. Moravcsik suggests that publications

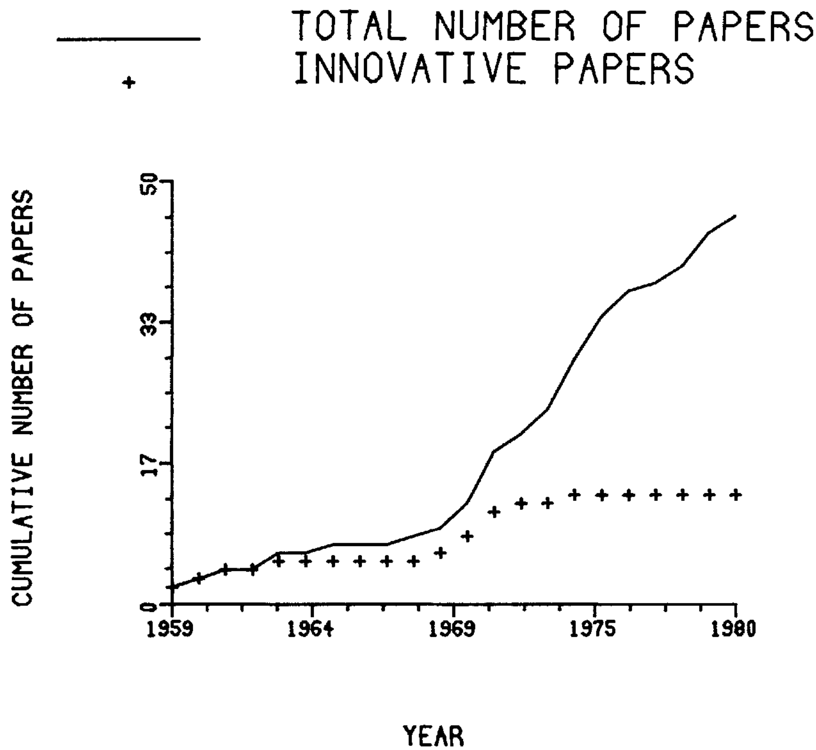

Fig. 7. Numbers of innovative papers and total papers published on obsolescence, 1944-80.

TABLE 7
CITATIONS PER ARTICLE FOR PAPERS ON OBSOLESCENCE, 1944-77

Article Category	*No. Papers*	*Average No. Citations*	*Minimum No. Citations*	*Maximum No. Citations*
New ideas and theory	13	12	1	28
Applications	11	7	0	14
Reviews	3	6	4	8
Other	19	4	0	23

and citation counts may be good first approximations to a measure of scientific growth: "The task then is to estimate the size of the correction to this approximation and to construct more refined but equally practical versions of these measures which take into account these corrections."[29]

Chubin and Studer have similar reservations about the use of citations as indicators of importance or innovation. In a study of 656 articles about research on a DNA polymerase "reverse transcriptase," they noted that "only the force of facts (e.g., Baltimore and Temin and Mixutani *did* independently discover the DNA polymerase) keeps the larger, well-funded laboratories of Spiegelman and the National Cancer Institute from swamping the citation structure."[30] Chubin and Moitra classify citations as essential (basic and subsidiary), supplementary (additional and perfunctory), and negative (partial and total). In a study of 443 references in forty-three articles in high-energy physics, they found 57.1 percent of the citations were either supplementary or negative.[31]

Forecasts

In 1963, Price said: "There is a possibility the exponential law is breaking down."[32] Exponential growth cannot go on forever. Recent figures seem to indicate that this change is indeed occurring. Price predicts that, when limits to growth are imposed on such a process, there will be various reactions: escalation of a new process, loss of definition of the old process, divergent (i.e., widely fluctuating) oscillations, or oscillations converging to the limit. Like Moravcsik, he feels changing communication patterns among scientists, brought about by new technology, will lead to a situation in which publications are of secondary value in communicating innovations—for popularization rather than research needs.

Rescher believes that this "quality drag" principle—i.e., that exponential increase in the total number of papers is needed to produce a linear increase in the number of first-rate papers—means that, eventually, the pace of innovation (i.e., first-rate findings) will begin to decline.[33] He regards the exponential increase in publication not as useless verbiage but as the useful and necessary inputs needed for genuine advances. However, in an age of dwindling resources, the world can no longer afford exponential input. Thus, growth in number of publications will become linear—perhaps has already become linear in the seventies. The growth in cumulative number of first-rate publications will then be logarithmic, i.e.,

$$F_o(t) = \log_e(a+bt),$$

and the continuous growth rate will become

$$b/(a+bt).$$

In other words, the further into the future we go, the fewer the additional number of first-rate publications. We are moving from an exponential growth past to a linear growth future.

To conclude, many papers have tried to estimate the growth of knowledge in various ways, and as many questions have been raised about the validity and reliability of bibliometric measures for this process. It appears that, for the "growth of knowledge" subfield, the time is not yet ripe for a logarithmic decline in the number of first-rate papers. There is an obvious need for better compilations of statistics on numbers of publications in the various disciplines on a worldwide scale, for informed, critical assessments of the amount of new knowledge contributed by these publications, and for enhancements and refinements of the present bibliometric techniques (citation and publication counts), so that valid measures of knowledge growth may be obtained. Also, studies of literature growth need to become more exact in the description of their models and more rigorous in the application of statistical tests to determine how well these models fit reality. Only then will bibliometrics be able to provide accurate, useful descriptions and predictions of knowledge growth.

References

1. A. Conan Doyle. Quoted *in* Nicholas Rescher. *Scientific Progress*. Pittsburgh: University of Pittsburgh Press, 1978, p. 54.

2. Popper, Karl. *Objective Knowledge; An Evolutionary Approach*. Oxford: Clarendon Press, 1972, p. 144.

3. Rescher, *Scientific Progress*, p. 48.

4. Price, Derek de Solla. *Little Science, Big Science*. New York: Columbia University Press, 1963; and ________ . *Science Since Babylon*. New Haven, Conn.: Yale University Press, 1961.

5. Crane, Diana. *Invisible Colleges*. Chicago: University of Chicago Press, 1972.

6. Lawson, J., et al. "A Bibliometric Study on a New Subject Field; Energy Analysis." *Scientometrics* 2(1980):227-37.

7. Frame, J. Davidson, et al. "An Information Approach to Examining Developments in an Energy Technology: Coal Gasification." *Journal of the ASIS* 30(July 1979): 193-201.

8. Crane, *Invisible Colleges*; Sullivan, Daniel, et al. "The State of Science: Indicators in the Specialty of Weak Interactions." *Social Studies of Science* 7(May 1977):167-200; and Menard, Henry W. *Science: Growth and Change*. Cambridge, Mass.: Harvard University Press, 1971.

9. Braun, T., et al. "An Analytical Look at Chemical Publications." *Analytical Chemistry* 52(May 1980):617A-29A.

10. Bottle, R.T., and Rees, M.T. "Liquid Crystal Literature." *Journal of Information Science* 1(May 1979):117-19.

11. Goffman, William. "Mathematical Approach to the Spread of Scientific Ideas—the History of Mast Cell Research." *Nature* 212(29 Oct. 1966):449-52.

12. ———, and Warren, Kenneth S. "The Ecology of the Medical Literatures." *American Journal of the Medical Sciences* 263(April 1972):267-73.

13. ———. "A Mathematical Model for Analyzing the Growth of a Scientific Discipline." *Journal of the ACM* 18(1971):172-85.

14. Bennion, Bruce, and Neuton, Laurence. "The Epidemiology of Research on 'Anomalous Water.' " *Journal of the ASIS* 27(Jan.-Feb. 1976):53-56.

15. Small, Henry G. "A Co-Citation Model of a Scientific Specialty." *Social Studies of Science* 7(May 1977):139-66.

16. Gilbert, G.N. "Measuring the Growth of Science: A Review of Indicators of Scientific Growth." *Scientometrics* 1(1978):9-34.

17. Moravcsik, Michael J. "Measures of Scientific Growth." *Research Policy* 2(Oct. 1973):266-75.

18. May, Kenneth O. "Quantitative Growth of the Mathematical Literature." *Science* 154(30 Dec. 1966):1672-73.

19. Oliver, F.R. "Methods of Estimating the Logistic Growth Function." *Applied Statistics* 13(1964):57-66.

20. Baker, Dale. "Recent Trends in the Growth of the Chemical Literature." *Chemical and Engineering News* 54(1976):23-27.

21. Carpenter, M.P., and Narin, F. "The Subject Composition of the World's Scientific Literature." *Scientometrics* 2(1980):53-63.

22. Hulme, Edward W. *Statistical Bibliography in Relation to the Growth of Modern Civilization*. London: Butler and Tanner, 1923.

23. Gottschalk, Charles M., and Desmond, Winifred F. "Worldwide Census of Scientific and Technical Serials." *American Documentation* 14(July 1963):188-94.

24. Bourne, Charles P. "The World's Technical Journal Literature." *American Documentation* 13(April 1962):159-68.

25. May, Kenneth O. "Growth and Quality of the Mathematical Literature." *ISIS* 59(Winter 1968):363-71.

26. Ibid., p. 368.

27. Gosnell, Charles F. "Obsolescence of Books in College Libraries." *College & Research Libraries* 5(March 1944):115-25.

28. Moravcsik, "Measures."

29. Ibid., p. 275.

30. Chubin, Daryl E., and Studer, K.E. "Knowledge and Structures of Scientific Growth." *Scientometrics* 1(1979):185.

31. Chubin, Daryl E., and Moitra, Soumyo D. "Content Analysis of References." *Social Studies of Science* 5(Nov. 1975):423-41.

32. Price, *Little Science*, p.19.

33. Rescher, *Scientific Progress*.

Appendix

Statistics Used for Graphs in the Text

The counts upon which the figures are based are as follows:

Figure 1

Year	*Chemical Abstracts*	*Year*	*Chemical Abstracts*
1907	11,847	1944	43,700
1908	15,169	1945	33,672
1909	15,459	1946	39,578
1910	17,545	1947	39,288
1911	21,682	1948	43,996
1912	23,194	1949	53,441
1913	26,630	1950	59,098
1914	25,115	1951	63,033
1915	18,981	1952	70,147
1916	16,108	1953	75,091
1917	15,945	1954	80,615
1918	13,881	1955	86,322
1919	15,240	1956	92,396
1920	19,326	1957	102,525
1921	20,451	1958	118,930
1922	24,098	1959	127,196
1923	25,315	1960	134,255
1924	26,643	1961	146,893
1925	27,097	1962	169,351
1926	30,238	1963	171,404
1927	33,491	1964	189,993
1928	39,135	1965	197,083
1929	48,293	1966	220,303
1930	55,146	1967	242,527
1931	52,728	1968	232,508
1932	59,461	1969	252,320
1933	66,153	1970	276,674
1934	61,570	1971	308,976
1935	63,413	1972	334,426
1936	64,572	1973	321,005
1937	64,735	1974	333,642
1938	66,928	1975	392,234
1939	67,108	1976	390,905
1940	53,680	1977	410,137
1941	50,494	1978	428,342
1942	45,646	1979	436,887
1943	43,669		

Figure 2

Year	Number of Abstracts		
	Science Abstracts	Biological Abstracts	Chemical Abstracts
1960	21,410	72,530	134,255
1961	21,160	87,000	146,893
1962	24,240	100,790	169,351
1963	26,000	75,710	171,404
1964	31,000	107,100	189,993
1965	34,000	110,120	197,083
1966	38,000	120,100	220,303
1967	40,790	125,030	242,527
1968	50,480	130,020	232,508
1969	49,610	135,010	252,320
1970	79,830	140,030	276,674
1971	84,340	140,020	308,976
1972	85,180	140,000	334,426
1973	81,350	140,040	321,005
1974	83,370	140,020	333,642
1975	87,630	140,020	392,234
1976	74,180	142,510	390,905
1977	91,670	145,010	410,137
1978	96,580	149,010	428,342
1979	101,240	154,990	436,887

Figure 3

Year	Number of Abstracts		
	Historical Abstracts	International Political Science Abstracts	Sociological Abstracts
1960	2,925	1,461,000	1,903
1961	2,776	1,510,000	2,322
1962	3,096	1,415,000	2,952
1963	3,926	1,355,000	3,810
1964	3,623	1,467,000	6,062
1965	3,363	1,471,000	4,262
1966	3,516	1,492,000	5,130
1967	3,527	1,574,000	5,434
1968	3,417	1,450,000	5,969
1969	4,180	1,693,000	6,019
1970	4,015	2,206,000	6,000
1971	6,406	2,244,000	6,981
1972	6,359	2,998,000	7,190
1973	7,607	4,555,000	6,689
1974	7,244	4,955,000	6,982
1975	8,779	5,015,000	7,687
1976	9,094	5,039,000	7,289
1977	15,414	5,040,000	8,267
1978	15,675	5,075,000	8,339
1979	15,692	5,105,000	0

Figures 4 and 5

Year	Number of Abstracts: Library and Information Science Abstracts	Psychological Abstracts
1960	1,003	8,532
1961	968	7,353
1962	986	7,700
1963	1,052	8,381
1964	1,054	10,500
1965	1,104	16,619
1966	1,106	13,622
1967	1,053	17,202
1968	1,226	19,586
1969	2,567	18,068
1970	2,858	21,722
1971	2,619	23,000
1972	3,177	17,976
1973	3,037	24,409
1974	3,837	25,558
1975	3,870	25,542
1976	3,781	24,687
1977	4,721	27,004
1978	4,886	26,292
1979	4,217	29,714

Figure 6

Year	No. of First Editions
1967	79,289
1968	78,875
1969	87,604
1970	95,433
1971	97,469
1972	103,679
1973	102,237
1974	100,958
1975	112,300
1976	110,715

Figure 7

Year	Number of Innovative Papers	Total
1944	1	1
1959	1	1
1960	1	1
1961	1	1
1963	1	2
1965	—	1
1968	—	1
1969	1	1
1970	2	3
1971	3	6
1972	1	2
1973	—	3
1974	1	6
1975	—	5
1976	—	3
1977	—	1
1978	—	2
1979	—	4
1980	—	2

This Page Intentionally Left Blank