Dollars-and-Cents Value of Efficient Presentation. A Plea for the Retention of the Well-Produced Primary Journal

R. SCHOENFELD

Editorial and Publications Service, CSIRO, East Melbourne, Victoria 3002, Australia

Received August 8, 1980

This paper attempts to demonstrate, with the help of a simple economic model, that the present trend toward low-cost and low-quality publication of scientific work may be beneficial for the individual publisher but harmful to the community as a whole.

The science of chemistry may be advancing with giant strides, but the quality of chemical publications is diminishing. Reading a chemical paper in a primary journal these days is not so much a stimulating experience as an obstacle race. The text bears all the hallmarks, or rather claw marks, of hasty and unloving editing; the layout is confusing; the print is hard to read. I am a professional editor nearing the end of my career, and it is very painful for me to watch the decay of the primary journal. There are some distinguished exceptions, such as *Angewandte Chemie*, but the general trend is inexorably downward from the standards of only two decades ago. One leading British journal, once a trend setter for terse, precise prose, has now gone overboard in its search for succinctness to such an extent that one no longer reads its papers, one decodes them. No longer do we admire, in the pages of *Helvetica Chimica Acta*, the superb hand-setting of chemical formulas that the master craftsmen of Birkhäuser Verlag delighted in when Ruzicka and Reichstein were at the height of their achievements. And in one commercial journal, of distinguished origins, the confusion of the layout is only matched by the illegibility of the type.

Superficial matters, one might object; the leading journals have simply, in a time of economic crisis, got rid of some unnecessary aesthetic frills, but the flow of science continues undisturbed. It will be argued in this paper, on strictly economic grounds, that this objection is not justified; inefficient presentation does retard the flow of science and in the end costs more than it saves. Moreover, the lack of superficial frills I have mentioned may be a symptom of a deeper disease. If a journal, to save money, leaves an author's badly drawn structural formulas unchanged, it may well, to save money, cut down on the time available for expert textual scrutiny during copy-editing, and it may even, to save money, oversimplify the refereeing process.

While the scientific publishing industry, under financial stress, has turned out inferior products during the past 20 years, the information retrieval industry, buoyed by increasing funds, has flourished. I certainly do not begrudge this success; it is very thrilling for me in Australia to be able to enter into almost-instant contact with a data base in California and conduct literature searches. But the present state of affairs is that ever more elaborate conveyor belts are being constructed to transport ever shoddier merchandise toward the user. Every keyword search or substructure search, no matter how ingeniously conducted, has as its output a set of literature citations. To these citations correspond published papers. And these papers have to be read.

Modern informatics has succeeded in offering the user a more complete menu, and to heap the mental food more quickly on his plate. But it has not altered the basic digestion process. It is time that we admitted this and considered the scientific publication process not from the narrow viewpoint of the publisher, author, or data archivist but from that of the entire community.

DATA BASE

The economic crisis beginning in the late 1960s and the steady advance of electronic innovations have caused much introspection in the scientific publishing industry (e.g., ref 1). The present paper relies for its hard data above all on the report prepared by King Research Inc. for the National Science Foundation.[2] This is a massive and painstaking attempt to quantify everything that is quantifiable in the scientific publishing industry. The King report suffers, perhaps, from being too all-embracing, and it may be that it gives "fringe" journals undue weight compared to "core" journals; some of its assumptions seem to me questionable and others are irreconcilable with my Australian experience. Overall, however, from this imposing array of quantitative data some indisputable near-quantitative truths emerge. The King report was written to compare the existing "paper-based" system of scientific publishing with the "Electronic Alternative". It found no present economic advantage for the latter but nevertheless considers it desirable. The present article takes a far more conservative stance; the fact that it relies on data obtained by holders of the opposite philosophy may protect it from a charge of bias.

The King report turns, as it were, a steady documentary camera on all the nooks and crannies of the publishing industry. Here, we shall attempt to pull the camera back far enough to include, in our picture frame, not just this specialist industry but the entire community. Occasionally, we shall also cut to closeups of the scientist at his desk and observe his reading habits.

SCIENCE MACHINE

The community at large views science as a black box into which it feeds problems and out of which it hopes to see solutions emerge. The benefits derived from this problem-solving process can generally be expressed in dollars and cents, but there are some (say the conservation of the environment or the exaltation of placing an astronaut on the moon) that cannot be so expressed. Still, modern economists have found ways around this difficulty (for instance, by substituting for the nonquantifiable benefits the value of those quantifiable benefits the community sacrifices to obtain the former). Thus we may say that when a community spends an amount S on science, it expects to gain from this amount benefits whose value B exceeds S; moreover the community will adopt the economic strategy of trying to obtain the maximum B for a given S. All this is, of course, elementary economics and

0095-2338/81/1621-0061$01.25/0

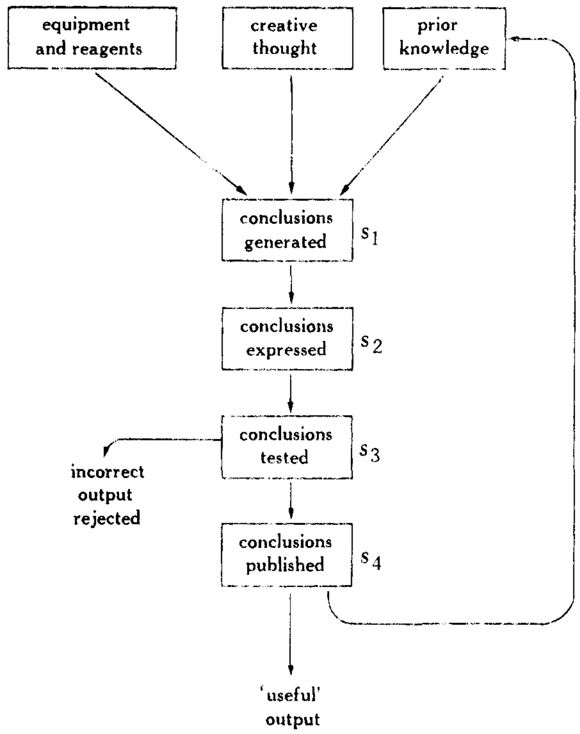

Figure 1. The science machine.

well-known; but the point should be made here that the community is exclusively interested in getting the highest value for B and does not care whether, in the pursuit of this goal, the amount S is spent on diffractometers, mainframes of retrieval systems, or printing presses.

Figure 1 offers a sketch, accurate enough for our purposes, of the contents of that black box, the "science machine". These comprise, above all, a reactor into which are fed prior knowledge (stored in libraries), instruments and reagents (kept in laboratories), and creative thought (emerging from the researcher's mind). The interaction of these three generates "conclusions", which are at that stage still stored in the researcher's mind and notebooks and thus not accessible to the community. These conclusions must thus be elaborated in a further reactor (i.e., the researcher's writing desk and secretarial office). At the end of this operation we have at last something tangible in the form of a manuscript, but again this is of doubtful benefit to the community because the conclusions may be partly or wholly wrong. Hence the conclusions have to enter a further reactor, in which they undergo essentially a purifying process (the purifying agent is the creative thought of referees and editors). If (and that is not certain) the conclusions survive this process, they are then fed into another reactor in which the publishing operation is carried out. (It would be wrong to consider this operation merely a packaging process; the conclusions remain under scrutiny during copy-editing and may undergo further, sometimes very considerable, purification.) The finished product, a paper or patent, is then available for public inspection. If it is directly useful to the community it becomes an input to the community's culture or technology; if it is not directly useful, it is recycled to become part of prior knowledge. The four "reaction steps" in Figure 1 have been labeled s_1, s_2, s_3, s_4; the sums spent on them will be called correspondingly S_1, S_2, S_3, S_4.

"INSTITUTIONS" AND "PUBLISHING HOUSES"

Steps s_1 and s_2 take place in institutions (industrial research laboratories, universities, government-supported laboratories, etc.), and step s_4 is carried out in publishing houses. As for step s_3, it begins in a publishing house (or an extension thereof,

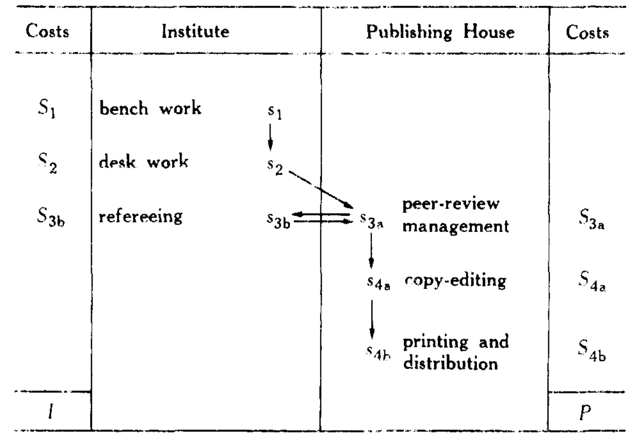

Figure 2. A list of all the communication steps discussed in this paper and their costs.

such as the office of an academic who has agreed to act as editor), moves back into an institution (where the referee prepares his report), and is finished off in the publishing house again (where the editor conducts all necessary correspondence and mediates between author and referee). The amount of money S_3 necessary to conduct this operation has thus two components: the sum S_{3a} spent by the publishing house and that, S_{3b}, contributed by the institution (which almost invariably "donates" the referee's time). If we define the expense incurred by the institutions in doing research as I, we have

$$I = S_1 + S_2 + S_{3b}$$

In the expense incurred by the publishing houses—let us call it P—we have already identified the term S_{3a}, the "peer-review function". The costs of step s_4, during which the accepted manuscript is transformed into the published article, will also be divided into two components. The first of these, S_{4a}, pays for "copy-editing"—those operations for which specialized editorial staff is required. The second, S_{4b}, accounts for typesetting, printing and mailing operations as well as the cost of paper, ink, machinery and so forth. Thus we have

$$P = S_{3a} + S_{4a} + S_{4b}$$

All operations and their costs are identified in Figure 2.

ECONOMY OF THE INSTITUTIONS

Consider an institution entirely dedicated to research and its publication. Over any given period, the addition an individual scientist employed by this institution causes to its expenses can be expressed as the product np, where p is the scientist's pay over the period and n is a multiplier expressing the rate at which he uses up the institution's resources (such as equipment and the time of his support staff). As the scientist moves from his laboratory to his library or writing desk, p remains constant but n undergoes very significant changes. My observations in Australia show that where n is 4 during laboratory work it is likely to be 1.3 during reading and writing periods.

This has to be borne in mind as we now consider the way in which, according to the King report, the average "author" (i.e., a physical scientist whose aim is to do research and write papers) divides his time. Such a man works 48 h a week for (my guess) 49 weeks a year, hence 2350 h per annum. He writes one paper a year (actually two papers, each with one coauthor) and the time spent writing (task s_2) is about 90 h (3.8%). Further he spends 20 h (0.9%) refereeing other author's papers (task s_{3b}).

These data concerning an idealized institution can be applied to the real world only with much caution, because there the "author" is apt to spend much time teaching and may be getting free student labor in return. I have, however, been able to confirm the above data, in a rough way, in Australia, where a sizable proportion of the scientist population works in government-funded research laboratories. The only change to the data I can suggest is that the average author writes 1.5, rather than 1, papers per annum.

This would bring the time spent on s_2 to 5.7%. However, the action of the multipliers *n*, described above, ensures that the cost S_2 is no higher than 2% of the total *S*, with S_1 being 98%; S_{3b} is negligible in terms of *S*. (It is not negligible in terms of the much smaller sum *P*, but this need not concern us here, seeing that S_{3b} is donated by the institutions.)

ECONOMY OF THE PUBLISHING HOUSES

Among the three terms that make up *P*, the "printing and mailing" costs S_{4b} predominate. Among commercial publishers it is a common rule of thumb that the "scientific" expenses S_{3a} and S_{4a} should be kept to no more than 20%; i.e., $S_{4b} \geq 0.8P$. In journals published by learned societies, the "scientific" expenses tend toward a greater proportion of the total. The data of the King report[2] (Annex II, Tables 44–46) lead to an average value of $S_{4b} = 0.73P$ for physical science journals in the United States in 1975.

Unfortunately, the King report does not differentiate between S_{3a} and S_{4a}. The American Chemical Society, however, has conducted a particularly conscientious analysis of its publishing budget, and very interesting studies have recently been published by Bowen[3] and Brogan.[4] The first of these deals with the cost S_{3a} of the peer review function. For the best-known primary journals of the ACS this varies from 8% to 15% of *P*. Brogan's paper[4] analyzes the cost of copy-editing; S_{4a} (the definitions may not completely overlap but are certainly very close) and arrives at a value of 3–9% of *P*. The "most mathematical" and "most illustrated" primary journals would obviously be near the top of this range.

As a handy approximation, we may thus assume that S_{3a} is 15%, S_{4a} 10%, and S_{4b} 75% of *P*. The last-named term is too well established to be subject to major error, and even wild fluctuations of the other two terms will not affect the conclusions about to be drawn.

Having thus obtained an estimate of the proportions of the three cost factors, we are now able to evaluate the strategies used by publishers to reduce the overall cost *P*.

ATTEMPTS TO REDUCE THE "PRINTING AND MAILING" COSTS S_{4b}

The obvious place to look for reductions in the cost of the publishing process is in the S_{4b} (printing and mailing) area, and it is here that the greatest changes have been made. Detailed discussion of composition and printing technology is beyond the scope of this article; the changes will be considered here only insofar as they affect the quality of the presentation of the text.

There have been attempts to reduce the cost of paper (and, within limits, the cost of postage) by shrinking the type size and using condensed typefaces, so as to be able to print more characters per page. Readers will readily call to mind various journals in which the perusal of even a short article causes severe eyestrain. I believe that this type of economy should be firmly condemned; most journals have only moderate press runs (1000–5000 copies) and thus the price of paper is not a major component in S_{4b}.

Worthy of far more serious consideration are the changes that have taken place in the composition process. Twenty years ago all major journals were printed by hot-metal typesetting; this process is now rapidly being supplanted and will certainly not survive the present decade. Its dominant successor is computer-assisted photocomposition. This is capable of yielding (and already does yield, in the hands of certain publishing firms) results comparable to those obtained by the older method and at much reduced cost. This change is thus to be welcomed, but there are certain dangers and difficulties associated with the computer-assisted photocomposition of scientific texts that deserve identification and discussion.

There is, first of all, the threat of the disappearance of the skilled compositors, who, in former times, were the editors' invaluable collaborators. Such craftsmen were familiar with routine chemical terminology and with the aesthetics of tabular setting; they could perform their task with a minimum of typographical guidance from the editors. Their place is now being taken by, in Brogan's[4] words, "a highly mobile staff of data-entry operators with little or no composition experience". This means that the copy editor, whose mind should be on the chemistry of the paper he edits, has to interrupt his deliberations to insert a profusion of typographical symbols. He will thus either take longer (so that the decrease in S_{4b} is partly offset by an increase in S_{4a}) or do an inferior job.

Next comes the problem of "merged material". Various publishers embraced computer printing in such haste that they did not await the arrival of the appropriate software for the setting of nonroutine text such as mathematical equations, figure captions, and tables. These items had to be set by a different method (generally "typewriter composition") and then to be inserted into "windows" left in the computer-set text. This procedure was a readymade source of errors, confusion among symbols, and aesthetic infelicities. The required software is now readily available, but the trend toward lower standards, especially in the setting of tables, will be difficult to reverse.

Another consequence of the switch to computer composition (unfavorable in my opinion) was the transformation of the footnote into a "tailnote". Placing a footnote at the bottom of its proper page required considerable skill of a human compositor; for a while the task was too difficult for the computer. The advent of computer setting thus meant that all such notes were banished to the end of the paper, so that the reader's concentration was forever upset by the necessity of commuting between the "current" and the "tail" page. It is a source of great satisfaction to me that "footnote capability" is now becoming available and that some American computer-set journals have restored the footnote to its proper place.[5]

So much for computer setting. Another widespread strategy for reducing production costs is to dispense with the typesetting process altogether. Here, the publisher simply asks the author to produce a typescript in conformity with the editorial rules of the journal and having a specified format; this typescript is then photographed and reproduced by a photolitho process. The product is, inevitably, much less attractive and much harder to read than printed text; but the publisher is able to congratulate himself on having effected an important saving in the S_{4b} area. However, this publisher's gain is the author's loss; the publisher has simply transferred to the author the expense of producing the final, error-free version of his paper. In other words, costs have been transferred from S_{4b} to S_2 (or *P* to *I*) rather than saved.

Besides his saving on the S_{4b} budget the publisher may yet achieve another gain, namely the elimination of the copy-editing cost S_{4a}. Whenever this happens, the published paper loses the benefit of the copy-editing process. If the publisher is, however, conscientious, then the cost of copy-editing in the "typewritten" journal may be higher than in the "printed" one. For the former, the copy editor must mark his corrections on the submitted manuscript, and must then communicate his instructions for revision not to an experienced printer but to an author inexperienced in such matters, who will then pass them on to his inexperienced secretary.

Finally, the typewritten journal imposes a severe handicap on the contributor whose native language is not that of the journal, especially if he belongs to an underdeveloped community and does not have good secretarial services and modern typewriters available.

ATTEMPTS TO REDUCE THE COPY-EDITING COSTS S_{4a}

The advent of computer setting has made possible some perfectly legitimate, but modest, economies in this area. The task of shaping the text into page form, customarily performed in the past in the editorial office by "pasting up" galley proofs in the desired format, can now be carried out by the computer, and there exist some programs that assign each "inserted" item such as figures and tables its most favorable place. Such programs can be considered genuine advances, but only if they allow the editor authority to override the computer's judgment.

Unfortunately, these minor attempts at economy are completely overshadowed, at the moment, by another money-saving process: the advent of the "typewritten" journal. As mentioned earlier, some of these journals (mostly those published by learned societies) attempt to make their authors conform to a scholarly "house style". The great majority, however, rely on their editor (generally a busy academic performing the function in his spare time) to make some admonitory remarks concerning bad presentation and then print the author's revised version warts and all. The overall economic disadvantages of this procedure will be outlined in the next section.

CONSEQUENCES OF COST CUTTING: COPY-EDITING

I define copy-editing as all those changes to the chemical information and the presentation of a paper that do not form part of the peer review. Some publishers make a sharp distinction between the editor, who supervises the refereeing process and the mediation that often follows it, and the copy editor, who generally works in a distant office and only sees the manuscript after it has been accepted. In other journals both functions are performed in the same office and may even be embodied in the same person. But, even in the case of journals with two-tier structure, the functions cannot be entirely kept apart. A copy editor may notice something both referee and editor missed, and an editor may make some helpful remark (on semantics, nomenclature, or style) that falls within the realm of copy-editing.

The value of a copy-editor's contribution may be negative as well as positive; he may introduce errors as well as remove them. One of the advantages claimed for the typewritten journal is the certainty that the author's numerical values are protected from editors' and printers' errors and will survive the reproduction process unchanged. In spite of such arguments, however, the vast majority of authors would concur that the contribution of the copy editor is positive and substantial. Copy-editing at its best is a highly creative occupation and may improve the published paper very considerably.

In some communities, copy editors are the experts on chemical nomenclature; authors turn to them for advice as they would turn to a spectroscopical specialist, an analyst, or a statistician. Nearly everywhere, copy editors are the leading experts on the proper use of symbols and chemical terminology. Their intervention prevents the message of a chemical paper from being obscured by noise.

Of perhaps greater importance still is the copy editor's ability to act as a referee of last resort. Technical editing involves the scrutiny of every letter on every line, and thus very often brings to light nontrivial errors the referees skimmed over. In our Australian practice, contacts (by letter or telephone) between copy editors and authors are common. Cases have occurred where an author withdrew his paper as a result of a copy-editor's discovery of a mathematical error; it is not uncommon for authors, as a result of a communication from the copy editor, to return to the workbench for additional experiments.

Finally, copy editors are the guardian angels of those authors who find it necessary to submit papers in a language not their own.

It is for such reasons that any reduction in the cost S_{4a} may seriously endanger the overall operation depicted in Figure 1. The benefits of the copy-editing step are difficult to quantify. However, if one makes the assumption (with which the scientific community at large would concur) that good copy-editing raises the value of a paper (i.e., its impact on the reader) by 1%, then the money spent on this operation may seem a burden to the publisher but is undoubtedly a good investment for the community.

CONSEQUENCES OF COST CUTTING: LEGIBILITY

There is no doubt that we all prefer to read scientific text in a typewritten rather than handwritten form. Likewise, nearly all of us prefer a clearly printed version to a typewritten one, and for the same reason: word outlines become easier to recognize. Matter that is difficult to read causes us the kind of discomfort we experience when listening to a lecturer with a speech defect or having a heavy accent.

Psychologists (see Tinker[6]) have succeeded in quantifying the loss in reading speed we experience when passing from standard-size to small-size print and from justified to unjustified type—but only for brief and simple text. Even for such reading matter, the descent from the well-produced to the badly produced journal might entail a loss of reading speed of 5%, and this alone would reduce the efficiency of the average scientist by just under 1% (see Annex II, Table 100 of ref 2).

Economically significant though these losses in efficiency may be, they do not represent the main danger to the functioning of the science machine. The most insidious peril represented by the badly produced journal is not the eyestrain it causes to the *interested* reader, but the way it repels the casual browser. "Browsing" is the act of gliding one's eye, at several times reading speed, over a column of text. In a well-produced journal the eye, like a fisherman's net, will present the mind with a sufficient "catch" of words to allow the gist of the text to be guessed at. With bad type the catch is too low, and the mind wearies of the task.

The decisive importance of browsing to the proper working of the science machine has, in the recent past, often been overlooked. The success of the informatics industry has greatly reduced the degree of randomness in the average scientist's reading. This change is beneficial only *if it is not allowed to go too far.* A certain stochastic element is necessary in the choice of the material that reaches the research worker's in-tray. In its absence, we all degenerate to the status of caged hens in egg-laying factories, meekly digesting the fodder dished out to us by the "research profile" assigned by our computer.[7]

The reader will readily call to mind many great advances in science that occurred because a theorist encountered, during random reading, the results accumulated by an experimentalist. Popular imagination is aware of this random element in science and has turned it into the legend of Newton's apple. In fact, whenever we enter our library, we do so determined to shake a few apple trees, in the hope of beneficent fallout. The amount of time we spend in this activity has been estimated by the King report. Table 100 of Annex II shows that 18.4% of the articles read by physical scientists were "discovered" rather than deliberately searched for; 60.8% were read as the result of a deliberate search; for the remaining 20.8% the interviewees did not remember.

Thus, the scientist's right to take a random walk through his library is fundamental and must be defended—against delimiting data bases as well as parsimonious publishers.

OVERALL VIEW

The central fact on which our investigations now converge is that the cost P of publishing the research is extremely small compared to the cost I of doing the research. The King report does not offer a figure for the physical sciences alone, but does offer an overall value for the whole area of scientific publishing. In 1975, the costs P in this area were (ref 2, Annex II, Table 115) U.S. \$400 million. Research and development funds for that year were (ref 2, main volume, p 19) U.S. \$35 billion; hence

$$P = 0.011I.$$

In Australia, it is possible to obtain data for the physical sciences alone. As mentioned earlier, much of the research work is carried out in government-funded laboratories, and thus the cost of "doing a piece of research" can be simply obtained by dividing the average budget of such an institution by the number of papers (or patents) it publishes, and then comparing this to the appropriate publication costs. I find, working with crude budget figures, that P is about 1.7% of I, but this drops to 1.3% if various "hidden charges" are considered (such as pension plans).

From an Australian analogy, I can also offer a guess at the U.S. figure for physical sciences. Mention was made earlier of a multiplier n, with which scientists' salaries must be multiplied to arrive at the total expenditure of the institution. In Australia the average value of n is close to 3.3 (this value neglects the "hidden charges") and if this is applied to the data of the King report it leads to P being about 1.5% of I.

In what follows we shall, to be conservative, assume P to be 2% of I. Then, an "average" research project that cost an institute \$100 000 to complete (\$98 000 in the laboratory and \$2000 at the writing desk) will cost \$2000 to publish. Of these, \$300 will be spent on the conduct of the peer-review function (and it may well be that at this stage the community receives the best value for its research dollar). The cost S_{4a} of copy editing will be \$200, and \$1500 will go toward typesetting and distribution. Seen in this context the possible savings (at most \$200 on S_{4a} and \$400 on S_{4b}) seem almost derisory, even before their consequences are considered. Such penny pinching calls the value of doing research into question: a community that cannot afford a halfpennyworth of tar should not build ships.

But the consequences of the savings do have to be considered, and they become very obvious from Figure 1. The performance of the science machine depends on the smooth flow of pure material into the "prior knowledge" reservoir and hence into the s_1 reactor. The above saving in copy-editing may lead to a loss of \$1000 (it could easily be \$5000) in the value of the research. A paper so printed as to induce eyestrain may only be read halfway through by half its readers, and this may well mean a loss of \$25 000 to the community. A paper that repels browsers may be worth \$30 000 less than one that invites them.

CONCLUSION

The acquisition of prior knowledge consists of two steps: (i) the placing of the relevant information before the scientist and (ii) its absorption into the scientist's mind. In the exuberance over progress made in step i, step ii is apt to be forgotten; but this is unjustifiable. Reading a scientific paper is a far more complex intellectual task than reading a pulp-paper novel: information must not only be accepted but also constantly correlated with, and tested against, information already in the reader's possession and then stored in a permanent memory. If further difficulties are added to this already difficult task, it cannot be performed successfully, and the community is the poorer for this failure.

There exist cynical laymen, and perhaps cynical publishers, who believe that authors publish only to gratify their vanity and to achieve tenure and that no trouble need be taken over text no one will read anyway. But this belief is contradicted by the data base. The average physical-science paper published in the average journal accessible in the "average" U.S. scientific library has, it appears, 576 "uses" in the United States during its useful lifetime.[2] Converting "uses" of 17-min average duration into full perusals, and extrapolating from the U.S. to the whole scientific world, one arrives at a surprisingly high number of "reader equivalents" (my estimate is 150). Authors should feel a responsibility toward these readers; they should also consider that, by allowing their work to be butchered and disfigured, they appear to accept their publisher's low evaluation of it. Let us reverse the trend toward unscholarly publishing; if research is worth doing well then it is also worth publishing well.

APPENDIX: A NOTE ON THE ELECTRONIC ALTERNATIVE

The Electronic Alternative is a projected communications system which dispenses with print on paper. The paper is generated by one keyboard operation only, in the author's office; it is then telecommunicated to an editor and by him to a referee. After the shuttling back and forth that constitutes the peer review process has been completed by electronic means, the paper is electronically conveyed to a copy editor who no doubt makes his changes in front of a video terminal. The author proofreads by the same means and the finished product is, again electronically, fed into a central memory (for which the name National Periodicals System has been suggested[2]) as well as the individual information reservoirs of chosen (perhaps computer-chosen) colleagues. Throughout the process, no paper is needed, except perhaps when the ultimate user tires of his visual display terminal and calls for a printout.

The King report was written to explore the feasibility of such an alternative. The authors, being experts in electronic communication, do not hide their commitment in its favor; but their report very fairly states that, although the system is fully feasible, it does not (in the present state of the art) offer any economic advantages. The actual and projected systems come out about even. In forecasts of this nature, that is a bad sign; but the meticulous presentation of the data has certainly earned the authors the right to have the Alternative considered on its merits.

Our model (Figure 1) shows that the Electronic Alternative has one respectable merit and several grievous disadvantages. The merit lies in the speeding up of the publication process. Anything that speeds up the flow into the prior-knowledge reservoir certainly enhances the performance of the science machine. (It should be noted, however, that, always according to the King report, a great number of the author's colleagues, wherever they are geographically located, have advance information on the nature of the work before the paper is written.)

Among the many disadvantages of the Electronic Alternative three are particularly serious:

(i) We have discussed the importance of legibility. A video terminal or crude lineprinter at the user's end will not do; to be acceptable, the Alternative must be able to generate, at the outlet, copy as good as that prepared by the present top journals.

(ii) Overseas communities who could tap this reservoir only by expensive satellite transmission will be tragically disadvantaged. And will the system cope with accented letters and alphabets other than Roman? The English language has reached its present eminence by its remarkable suppleness and by the remarkable scientific performance of the anglophone communities. Let us take care lest leadership by excellence degenerate into dictatorship.

(iii) Most serious of all is the impediment the system offers to random reading. As I have stated earlier, science without browsing is not science.

I am no hidebound conservative. It may well be that the brilliant advocates of the Alternative will find ways to overcome these problems and still keep the system economic, and in that case I shall be just as happy as my neighbor to recycle my stacks of bound volumes. But let us not proclaim the millenium until we see the correct date on the calendar.

ACKNOWLEDGMENT

In preparing this polemic, I have had much help from concerned Australian colleagues, especially B. J. Walby and D. E. Boyd. (Boyd[8] has published a model similar to my own; it deals mainly with the signal-to-noise ratio of scientific communication.) J. Glover has supplied much helpful information on the financial structure of scientific institutions. I owe a special debt to Dr. M. Brogan, of the American Chemical Society, who first drew my attention to the King report and whose informed criticism demolished an earlier version of this paper.

REFERENCES AND NOTES

(1) "Economics of Scientific Publications", Council of Biology Editors, Washington, 1973.
(2) D. W. King and N. K. Roderer, "Systems Analysis of Scientific and Technical Communication in the United States. The Electronic Alternative to Communication through Paper-Based Journals", National Science Foundation, Washington, DC, 1978, with four annexes.
(3) D. H. M. Bowen, *Scholarly Publishing*, **11**, 43 (1979).
(4) M. Brogan, *Scholarly Publishing*, **11**, 47 (1979).
(5) Some authors with a penchant for rambling asides actually prefer tailnotes. It would be possible for a journal to distinguish between footnotes and tailnotes by numbering the first and labeling the latter with asterisks and daggers. The *Australian Journal of Chemistry* distinguishes "citation" and "information" footnotes in such a manner, but prints both at the foot of the page, so that the reader can assess their importance to him at a glance. The present tailnote has been inserted as a warning example, to show what annoyance such notes can cause.
(6) M. A. Tinker, "Legibility of Print", Iowa State University Press, Ames, 1963.
(7) R. Schoenfeld, *Chem. Aust.*, **44**, 206 (1977).
(8) D. E. Boyd, *Aust. Phys.*, **15**, 39 (1978).

Analysis of Keywords in Chemistry

SHIZUO FUJIWARA, MASANORI YOKOYAMA,* and SHUICHI UEDA†

Department of Chemistry, Faculty of Science, The University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

Received August 27, 1979; Revised Manuscript Received October 13, 1980

From the ~10 000 000 keywords in 3 volumes of *CA Condensates* for 1977 and 1978, 16 000 keywords were assembled from the 200 most frequently occurring ones in each of the 80 sections and analyzed for cross-correlation, characteristic features, and relationship to 10 000 terms in a separate Japanese Chemical Society (JCS) list.

Each abstract in *CA Condensates* is represented in condensed form by bibliographic data and keywords. The former consists of the title, author(s) name, and journal name with volume number, issue, page, and year of publication. A few keywords, representing the content of each article, convey to the reader the essence of the subject of the document. Accordingly, keywords profile the documents and thus each field of chemistry is described by its assembly of keywords.

Chemical Abstracts has 80 sections, each referring to a specific field of chemistry. According to Chemical Abstracts Service, the *CA* section arrangement has never been intended as a model for an overall classification scheme in chemistry but simply reflects the *CA* subject coverage based on the relative amount of information published in various general and specific fields. Hence, the section arrangement has been changed occasionally to reflect progress, changes, and trends occurring in the published world of chemistry. For instance, as applied chemistry and chemical engineering have developed over the years, more technology- and product-oriented *CA* sections, such as "Plastics Manufacture and Processing" and "Surface-Active Agents and Detergents", were created in contrast to the more traditional subdivisions of chemistry, such as "Inorganic Analytical Chemistry" and "Electrochemistry". Although we have "Subject Coverage and Arrangement of Abstracts by Sections in Chemical Abstracts" as a useful explanation for the classification, it still seems worthwhile to carry out an analysis of the *CA* classification.

Table I. Statistics of the Volumes of *CACon*

Vol.	DC	KWC	no. KW
86	199 309	3 140 549	348 304
87	210 632	3 327 357	365 428
88	202 524	3 210 321	358 542

All keywords which appeared in volumes 86 (Jan–June, 1977), 87 (July-Dec, 1977), and 88 (Jan–June, 1978) of *Chemical* Abstracts were collected and examined. The frequency of occurrence of keywords was counted for each section of *CA Condensates*, and the 200 most frequently occurring keywords were collected for each section. The total number of keywords in each section is referred to as the keyword count number (KWC), and the total number of unique keywords per document as the number of keywords (No. KW). The number of articles was noted and referred to as the document count (DC). Cross-correlation of each keyword among the 80 sections was carried out. Keywords were grouped according to whether they can be used generally or only in chemical documentation.

* To whom correspondence should be addressed at Fujitsu Corporation, Ohta-ku, Tokyo.
† Department of Library Sciences, Keio University, Mita, Minato-ku, Tokyo.