TECHNICAL INFORMATION FLOW PATTERNS

M. M. Kessler
Lincoln Laboratory, * Massachusetts Institute of Technology

Summary

A study of the bibliographies of a large number of articles in physics and electrical engineering indicates that definite patterns exist for the flow of technical information. Quantitative data are presented on the flow of information between countries, between cultural and functional groups, and between past and present. An analysis of the numerical data indicates that these flow patterns are deeply rooted in the dynamics and evolution of scientific thought and engineering development. The analysis also discloses that extreme asymmetry exists between journals in their capacity as carriers of scientific information.

Numerical Data

Communication across Political Boundaries

A number of journals were analyzed for purpose of obtaining a rough measure of the flow of scientific information across national boundaries. The bibliographies in the indicated journals were sorted on the basis of country of origin. Table I shows the results for the January 1957 issue of the Physical Review.

<table>
<thead>
<tr>
<th>Reference to</th>
<th>No. of References</th>
<th>% of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Review</td>
<td>994</td>
<td>48.0</td>
</tr>
<tr>
<td>Other American</td>
<td>558</td>
<td>27.0</td>
</tr>
<tr>
<td>British</td>
<td>198</td>
<td>9.5</td>
</tr>
<tr>
<td>European</td>
<td>258</td>
<td>12.4</td>
</tr>
<tr>
<td>Russian</td>
<td>28</td>
<td>1.4</td>
</tr>
<tr>
<td>All others</td>
<td>33</td>
<td>1.6</td>
</tr>
</tbody>
</table>

We see that roughly half of the references in the Physical Review are to papers published in the same journal. Three quarters of all the references are to American journals. The remaining 25% of the references are distributed among a variety of European journals, mainly British. Note in particular the vanishing call on Russian reference material (1.4%).

* Operated with support from the U. S. Army, Navy and Air Force
We now consider the same process as it operates on Russian physicists. Table II gives the results of a count on the June and October 1957 issues of the Journal of Theoretical and Experimental Physics.

Table II
Geographic Distribution of References in the Russian Journal of Theoretical and Experimental Physics (June and October 1957)

<table>
<thead>
<tr>
<th>Reference to</th>
<th>No. of References</th>
<th>% of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>JETP</td>
<td>102</td>
<td>15.4</td>
</tr>
<tr>
<td>Other Russian</td>
<td>201</td>
<td>30.5</td>
</tr>
<tr>
<td>Physical Review</td>
<td>148</td>
<td>22.4</td>
</tr>
<tr>
<td>Other American</td>
<td>57</td>
<td>8.7</td>
</tr>
<tr>
<td>British</td>
<td>63</td>
<td>9.5</td>
</tr>
<tr>
<td>European</td>
<td>66</td>
<td>10.1</td>
</tr>
<tr>
<td>All Others</td>
<td>22</td>
<td>3.3</td>
</tr>
</tbody>
</table>

Note that on its home grounds the Russian JETP actually runs second to the Physical Review. The Russian physicists depend on British and other European literature roughly to the same extent as do the Americans, but they do not have the strong partiality to their own chief journal nor to any combination of journals in their political group.

Similar counts were made on Nuovo Cimento and Physica, Italian and Dutch journals of physics. The results are shown in Tables III and IV.

Table III
Geographic Distribution of References in the Italian Nuovo Cimento (Jan. - June 1958)

<table>
<thead>
<tr>
<th>Reference to</th>
<th>No. of References</th>
<th>% of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuovo Cimento</td>
<td>344</td>
<td>15.7</td>
</tr>
<tr>
<td>Other Italian</td>
<td>38</td>
<td>1.7</td>
</tr>
<tr>
<td>Russian</td>
<td>84</td>
<td>3.9</td>
</tr>
<tr>
<td>Physical Review</td>
<td>771</td>
<td>35.6</td>
</tr>
<tr>
<td>Other American</td>
<td>331</td>
<td>15.4</td>
</tr>
<tr>
<td>British</td>
<td>223</td>
<td>10.5</td>
</tr>
<tr>
<td>European</td>
<td>245</td>
<td>11.2</td>
</tr>
<tr>
<td>Others</td>
<td>135</td>
<td>6.2</td>
</tr>
</tbody>
</table>

To extend this picture somewhat beyond pure physics, a count was made on the Journal of Applied Physics (JAP) and the Proceedings of the Institute of Radio Engineers (IRE). In the latter case, we picked January, June, and September 1957 as representative issues. A special issue of the IRE devoted to a symposium on transistor technology was treated separately. The results are shown in Table VI. The Phys. Rev. data are reproduced for comparison.

Table IV
Geographic Distribution of References in the Physica (1957)

<table>
<thead>
<tr>
<th>Reference to</th>
<th>No. of References</th>
<th>% of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physica</td>
<td>225</td>
<td>21.6</td>
</tr>
<tr>
<td>Other Netherlands</td>
<td>71</td>
<td>5.9</td>
</tr>
<tr>
<td>Russian</td>
<td>30</td>
<td>3.0</td>
</tr>
<tr>
<td>Physical Review</td>
<td>249</td>
<td>23.9</td>
</tr>
<tr>
<td>Other American</td>
<td>85</td>
<td>8.2</td>
</tr>
<tr>
<td>British</td>
<td>159</td>
<td>15.2</td>
</tr>
<tr>
<td>European</td>
<td>121</td>
<td>11.6</td>
</tr>
<tr>
<td>Others</td>
<td>108</td>
<td>10.4</td>
</tr>
</tbody>
</table>

In view of the political and cultural polarization of modern society into East and West groupings, it is interesting to present the data of Tables I to IV in such a way as to illustrate the flow of physics from East to West and vice versa. Table V groups all references to American and European journals and compares them with those to Russian and other "iron curtain" journals.

Table V
The Flow of Information along the East-West Political Axis

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Western Jnls.</td>
<td>96.9</td>
<td>90.1</td>
<td>86.4</td>
<td>50.7</td>
</tr>
<tr>
<td>Eastern Jnls.</td>
<td>1.4</td>
<td>3.9</td>
<td>3.0</td>
<td>45.9</td>
</tr>
<tr>
<td>All Others</td>
<td>1.6</td>
<td>6.2</td>
<td>10.4</td>
<td>3.3</td>
</tr>
</tbody>
</table>

To extend this picture somewhat beyond pure physics, a count was made on the Journal of Applied Physics (JAP) and the Proceedings of the Institute of Radio Engineers (IRE). In the latter case, we picked January, June, and September 1957 as representative issues. A special issue of the IRE devoted to a symposium on transistor technology was treated separately. The results are shown in Table VI. The Phys. Rev. data are reproduced for comparison.

Table VI
Geographic Distribution of References in Phys. Rev., JAP, and IRE

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S.A.</td>
<td>75.0</td>
<td>70.5</td>
<td>77.0</td>
<td>78.0</td>
</tr>
<tr>
<td>British</td>
<td>9.5</td>
<td>12.0</td>
<td>11.6</td>
<td>5.3</td>
</tr>
<tr>
<td>European</td>
<td>12.4</td>
<td>11.7</td>
<td>8.7</td>
<td>8.6</td>
</tr>
<tr>
<td>Russian</td>
<td>1.4</td>
<td>2.9</td>
<td>1.2</td>
<td>2.4</td>
</tr>
<tr>
<td>Others</td>
<td>1.6</td>
<td>2.9</td>
<td>1.4</td>
<td>1.9</td>
</tr>
</tbody>
</table>
The data suggest the following conclusions.

1. The Physical Review is truly a definitive journal for physicists. It commands overwhelming dominance over all other journals as a carrier of information between physicists of all lands.

2. American physics is the chief source of information, not only for other Americans but for the international community of physicists.

3. American workers in fields of applied physics (as typified by authors in JAP and IRE) find their literature needs overwhelmingly satisfied by American journals.

4. European physicists draw heavily on American literature. Their coupling to the Russian literature is not significantly greater than that of American physicists.

5. The Western world is virtually self-sufficient with regard to physics. The Russian cultural sphere on the other hand draws heavily on the West for its information.

A close comparison of the various tables suggests that these conclusions are valid even if we take into account the language barrier between East and West.

Flow of Information to an Applied Field

A significant special case of information flow and retrieval concerns communication across disciplinary boundaries. We suspect that a retrieval or communications scheme designed to process physics literature for the physicist is not the same as what is needed to process physics literature for the chemist, the engineer, or the biologist. A related problem, particularly important in the dynamics of applied research and development, is the flow of information from basic research scientists to production engineers. If the coupling is tight and information flows freely, one may expect a low lag time between basic discovery and application. A study of certain reference statistics indicates that definite patterns exist in this information circuit.

The June 1958 issue of the Proceedings of the IRE was chosen for study. This issue is a symposium of 22 papers (353 pages) devoted to transistor technology. Of the fifty authors, forty-one indicated affiliation with industry, eight with universities, and one with the government. Transistor technology is of great interest to industrial and defense workers and yet it is new enough to have rather simple and short routes to the underlying sciences. For these reasons it was thought that a detailed analysis of the transistor issue would be instructive. Table VII shows the distribution of references among journals.

### Table VII

<table>
<thead>
<tr>
<th>Journal Distribution of References in the Transistor Issue of IRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference to</td>
</tr>
<tr>
<td>---------------</td>
</tr>
<tr>
<td>Physical Review</td>
</tr>
<tr>
<td>Proc. IRE</td>
</tr>
<tr>
<td>J. Applied Physics</td>
</tr>
<tr>
<td>Bell System Tech. Jnl.</td>
</tr>
<tr>
<td>British</td>
</tr>
<tr>
<td>German</td>
</tr>
<tr>
<td>Other Foreign</td>
</tr>
<tr>
<td>Russian, etc.</td>
</tr>
<tr>
<td>Miscel. Jnls. (American)</td>
</tr>
</tbody>
</table>

Tables VI and VII suggest that the transistor issue of the IRE shows the typical American pattern of bibliographic distribution and does not differ much from a random issue of the IRE. This invariance applies only if we consider the geographic distribution of references. Both cases follow the American pattern; over 75% of the references are to American and over 90% to Western journals. But when we analyze in detail the American journals for the two cases, an entirely different picture emerges. Table VIII is a break-down of the references to American journals in the two samples of the IRE.

### Table VIII

<table>
<thead>
<tr>
<th>Detailed Analysis of IRE References to American Journals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Articles References to IRE</td>
</tr>
<tr>
<td>---------------------------------</td>
</tr>
<tr>
<td>Physical Review</td>
</tr>
<tr>
<td>JAP</td>
</tr>
<tr>
<td>Proc. IRE</td>
</tr>
<tr>
<td>Other American</td>
</tr>
</tbody>
</table>

Whereas the average issue of IRE refers to the Physical Review 7.3%, the special transistor issue has 25.2% of its references to the Physical Review. Thus we see that the bibliographic count is sensitive enough to measure the degree of coupling between science and technology.

The authors who contribute generally to the IRE have a lesser coupling to the Physical Review than those who are preoccupied with the new and rapidly developing field of transistors. This analysis of information coupling between science and technology can be continued another step. We see from Table VIII that 25% of the references in the transistor issue of the IRE are to authors of papers in the Physical Review. We now ask who are these authors? Do they differ as a class from the usual authors in the Physical Review? In other words, do the contributors to...
the IRE transistor issue make contact with a representative group of Physical Review authors, or is there some transitional group of physicists who serve as a bridge between the general population of physicists and the industrial group?

Table IX presents the institutional origins of authors who publish in the Physical Review, Journal of Applied Physics, and IRE. In all cases column A refers to all authors in a random issue of the journal. Column B refers to those authors in the journal who were cited as references in the transistor issue of IRE.

Table IX

<table>
<thead>
<tr>
<th>Affiliation</th>
<th>Phys. Rev.</th>
<th>JAP</th>
<th>IRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>University</td>
<td>66 16.4</td>
<td>40.3 31.0</td>
<td>25.8 14.1</td>
</tr>
<tr>
<td>Industry</td>
<td>9 74.0</td>
<td>31.4 61.0</td>
<td>44.0 76.8</td>
</tr>
<tr>
<td>Government</td>
<td>13 6.3</td>
<td>47.5 5.4</td>
<td>10.0 0.6</td>
</tr>
<tr>
<td>Foreign</td>
<td>9 3.4</td>
<td>9.0 2.7</td>
<td>20.5 8.5</td>
</tr>
<tr>
<td>Others</td>
<td>3</td>
<td>1.8</td>
<td>-</td>
</tr>
</tbody>
</table>

The data show that ordinarily 66% of Physical Review authors have university connections and only 9% are from industry. The sub-group referred to by IRE authors more than reverses this picture; only 16.4% are university affiliated and 74% are from industry. Similar trends are apparent in the other columns of Table IX. On the basis of these limited data, it seems reasonable to assume that the coupling between basic science and its applied technology is tighter for the newer technologies and that the coupling is made through an intermediate group of scientists who form an intellectual bridge between the university and industrial community.

Flow of Information from the Past

At any given time scientists draw heavily on the accumulated experience of the past. Indeed, one of the greatest assets of the journal-article mode of communication is that it conserves the record of the past in an orderly and chronological manner. This coupling to the literature of the past was studied by plotting the number of references as a function of time into the past. Thus a distribution curve of references was obtained with age as the independent variable. Graphs 1, 2, 3, and 4 show such distributions for four journals. The integrated curves are shown in Figure 5. Although the curves are orderly, their statistical base is rather limited and one should be careful with conclusions. One may speculate that vigorous and fast-growing fields will show a sharp early rise and level off rather soon. Another hypothesis, however, may suggest that a sharp early rise indicates a superficial scholastic approach and that a more basic acquaintance with the literature of science would extend the curve farther into the past. Both hypotheses may be true, namely, workers in fast-growing, competitive fields may have no time to search the literature and thus confine their sources to current material. It is at any rate clear that this phenomenon has to be understood and absorbed into a serious retrieval scheme.

The Detailed Bibliographic Structure of a Single Journal

As a final example of bibliographic statistics, we mention a very detailed study that we made of the references in the Physical Review. The study was made for other purposes and involved a complete recording on IBM cards of all the references in 26 volumes of the Physical Review. Some of the results are mentioned here because they relate to the subject of this paper. Excluding the unpublished and non-periodic literature, the authors in the volumes made 74,599 references to journal articles. Of this number 45,592 or 60% were to articles in the Physical Review. The next five most frequently used journals contributed another 13.8% of the references and the next twelve in order of frequency contributed 12.2%. Thus 18 journals accounted for 86% of all the references. The remaining 14% of the references were distributed among some 650 journals of which 240 were mentioned not more than once in all the 26 volumes and 420 were mentioned four times or less.

Another by-product of this study that is relevant to our discussion is the following. In spite of the strong definitive position of the Physical Review, it is nevertheless true that any given paper in the Physical Review has a very low probability of ever being used as a reference. Indeed, the largest single class of papers never appears in the reference literature at all. It is hard to assume that this large group of papers are never cited because they are worthless. Other reasons must be sought.

Discussion

The over-all impression left by the data may be summarized as follows:

a) If we consider the scientific paper as a message unit and the journal the message carrier, if we accept that inclusion of a paper in a published list of references indicates that the message found a relevant receiver, and if we regard the population of Physical Review authors as a representative group of physicists, then there is a massive asymmetry and an overwhelming inhomogeneity in the capacity of the many

* This work was done in collaboration with Mr. Frank Heart of the Lincoln Laboratory and will be published elsewhere.
hundreds of journals to serve as carriers of the scientific message. The imbalance may be due to language barriers, cultural and political isolation, reputation of the journal we have examined and its availability or to a combination of these. The inhomogeneity exists, whatever the reason, and in its most extreme form gives rise to the definitive journal, such as the Physical Review. The data raise important questions about the design philosophy of retrieval systems. In view of this inhomogeneity, should the retrieval process and the flow channels be the same for all carriers or should they take account of the carrier's capacity? If account is to be taken of the carrier's capacity, should the system be designed to further reinforce the strong and efficient carriers at the expense of the less efficient, thus reducing the noise, or should we on the contrary take the attitude that the efficient carriers need less attention than the weak, and therefore concentrate on the latter and raise their efficiency? Is the definitive journal a desirable phenomenon or does it in the long run inhibit the communication process?

How should we approach the complex problem of injecting the results of Russian research into the main stream of American physics. Such an injection is certainly desirable, but it is not clear that a massive translation program and wide distribution of the translated material is the best way of doing it. Another method might be to encourage a small number of practicing American physicists to learn the Russian language and depend on them as the instrument of injection. Considering that the money and effort available for this purpose are limited, one should not blandly accept either method.

A retrieval and communication system, unless its contribution is trivial, will have sufficient feedback to strengthen or weaken the various elements of the communication process that now exists. It is therefore important to understand these elements and to design our system with them in mind.

b) The previous section concerned communication within the relatively homogeneous group of physicists who are in the habit of publishing in the Physical Review. The problem of communication across field boundaries or between scientists and engineers is a somewhat different matter. The limited amount of data that we have collected must be considered as a sample that only indicates the complexity of the problem. It would seem that in this case there is no definitive journal. Furthermore, meaningful communication seems to involve a chain of intermediaries that form a bridge between pure science and applied technology. Should a system be designed to encourage traffic along this chain of bridges, or should it attempt to short circuit the chain? That this is not a purely academic problem is clear from the experience of the Chemical Abstracts, a major communicative link between physics and chemistry. The Chemical Abstracts attempts to short circuit any bridging mechanism and bring physics directly to each chemist by abstracting practically the entire physics literature. This, together with other examples of short circuitry, has so overloaded its own channel that Chemical Abstracts is becoming increasingly awkward and bulky. They could have chosen not to include the physics literature in its abstracts and to depend on the various hyphenated journals, such as the Jn. of Physical-Chemistry, Chemical-Physics, Colloid Chemistry, etc. etc., to act as injectors of physics into the main stream of the chemical literature. This would appear to loosen the coupling between physics and chemistry and increase the time necessary for information transfer. But the decrease in volume of traffic could well compensate for the looser coupling and produce a more efficient system. The phenomenon of coupling between fields is at any rate important enough to be considered in the design of a retrieval and communication system.

c) The flow of scientific information in time, past to present, or the useful half life of a scientific message unit is another important element of our problem. Our numbers indicate a useful half life of some five to ten years. A rigorous definition of this phenomenon is not easy to come by. But it is obvious that no system of retrieval and communication can long survive without some method of purging its message population from time to time. This is not just a matter of eliminating poor material in favor of newer and better results. It is a curious fact that even the masterpieces of scientific literature will in time become worthless except for historical reasons. This is a basic difference between the scientific and belletristic literature. It is inconceivable for a serious student of English literature, for example, not to have read Shakespeare, Milton and Scott. A serious student of physics, on the other hand, can safely ignore the original writings of Newton, Faraday and Maxwell. The removal of a scientific paper from the retrieval system should not depend on a value judgment. The correct criterion should be based on the degree to which the paper's information has been metabolized into the flow stream of science. We could perhaps decide that once a paper has appeared in the citation literature a given number of times, it need no longer be carried as an independent message unit. This too needs more study. At the other extreme we have the phenomenon of a large group of papers, perhaps the largest single group, that is never quoted in other people's bibliographies. If these papers do not enter the bibliographic literature within, say, five years of publication, they may become effectively lost to the literature. In view of the careful editing process, it is hard to believe that this large group of papers is useless or redundant. If they are useless or redundant, the publishers and editors could well afford to review their processing criteria. On the other hand, a
continuing review of the citation literature over a
traveling five-year period may reveal a group of
papers that warrant more detailed channeling.

Concluding Remarks

Science and technology are approaching a
crisis in communication. It would be a mistake
to define this crisis entirely in terms of retrieval
problems. Indeed, the break-down in communi­
cation between scientists must itself be evaluated
in terms of the maturing sociology of science.
Communication, after all, is not the only aspect
of science that is in crisis. There are problems
of technical manpower shortages, the increasing
cost of scientific research, the growing imbal­
ance between basic and applied research, and
many others that relate to the emergence of sci­
ence as a major instrument of national favor,
stature and propaganda. But even if we confine
our attention to the communication problems only,
we must still remember that no single channel or
mode is likely to solve all our needs. The com­
ponents must be evaluated in terms of their con­
tribution to the over-all system performance.
Many systems problems must be studied if a work­
solution is ever to be achieved.

a) No system can be designed in an
economic and social vacuum. Even the vital
functions of national defense are subject to re­
straints. We must estimate the social and eco­
nomic limits that will govern our system and
optimize its function within these limits. How can
limited resources best be apportioned between
the various segments of the scientific community?
What part of the resources shall be assigned to
retrieval and other question-answer functions as
opposed to directing the flow of information on
the basis of generalized need-to-know criteria?
What are the information needs within well-
defined fields like physics as opposed to the flow
of information across field boundaries?

b) If there is a critical failure of com­
munication now, how will we know when improve­
ment has taken place? What test procedures and
criteria of performance can we use to evaluate
the system? Unless a figure of merit can be as­
signed to the system as a whole, the contribution
of any given component is always in doubt. One
cannot test the performance of a system by
measuring each component separately in terms of
its own parameters. Experience with large multi-
component systems indicates that one cannot
generally arrive at a figure of merit analytically.
It is usually necessary to build a model and eval­
uate new components in terms of their effect on
the model.

c) To what extent should the system
operate only when interrogated and to what ex­
tent should it operate on the initiation of its
built-in logic?

d) Should the system be local, regional,
or national? To what extent can communication
media other than the journal article be exploited
(radio, television, newspapers, remote printers,
etc.)?

e) What is the probable cost of various
systems, both initial capital investment and
operating cost? Shall it be financially self-
supporting or shall it be subsidized? Stability of
financing is of particular significance because
the system must have long-range continuity in
order to be at all effective.

f) Finally, we wish to stress that the
problems of science communication are not pri­
marily equipment and hardware problems. Nor
are they primarily problems in indexing, abstract­
ing, or retrieval. The significant problems are
in the area of systems design and systems logic.
At this time there is no single organization whose
professional competence and involvement embraces
the entire spectrum of problems. Such an organi­
zation is needed if a reasonably successful solu­
tion is expected.

Reference

1. See for example the series of papers in
"Area I - The Collected Papers of the
International Conference on Scientific
Figure 1  Distribution of References by Age
Physical Review 1957
Figure 2  Distribution of References by Age
Russian JETP, February, March, April 1957
Figure 3  Distribution of References by Age
Physica 1957
Figure 4  Distribution of References by Age
Figure 5  Cumulative Distribution of References by Age