REFERENCES

(1) National Science Foundation, "*Current Research and Development in Scientific Documentation,*" No. 8, NSF-61-29, May 1961.
(2) Helen L. Brownson, "Research on Handling Scientific Information," *Science*, December 30, 1960, vol. 132, No. 3444 pages 1922-1931.
(3) W.T. Knox, "Administration of Technical Information Groups," *Industrial and Engineering Chemistry, 51*, no. 3, pp. 48A-61A. March (1959); this is a condensation of the Symposium at the 134th meeting of the American Chemical Society, Chicago, Ill.
(4) Robert E. Maizell, "Measuring the Effectiveness of Technical Library Performance," *IRE Transactions on Engineering Management*, Vol. *EM-7*, no. 2, 69-72 June (1960).
(5) Peter V. Norden, "On the Anatomy of Development Projects," *IRE Transactions on Engineering Management*, Vol. *EM-7*, no. 1, March (1960).
(6) Peter V. Norden, "The Study Committee for Research, Development, and Engineering (SCARDE): A Progress Report and an Invitation to Participate," *IRE Transactions on Engineering Management*, Vol. *EM-8*, no. 1, March (1961).
(7) Albert H. Rubenstein, "Problems in the Measurement of Interpersonal Communication in an Ongoing Situation," *Sociometry, 16*, no. 1, pp. 78-100 February (1953). This article contains a bibliography of more than 40 articles describing techniques for studying communication within organizations.
(8) Operations Research Group, Case Institute of Technology, "*An Operations Research Study of the Scientific Activity of Chemists,*" November 1958, 69 pp. plus appendices. Processed.
(9) David B. Hertz and Albert H. Rubenstein, "Field Studies of Industrial Research and Development Groups," from "*Team Research,*" Department of Industrial Engineering, Columbia University, June, 1953.
(10) Albert H. Rubenstein and Robert W. Avery, "Idea Flow in Research and Development," "*Proceedings of the National Electronics Conference,*" 1959.
(11) Ralph R. Shaw, "*Pilot Study on the Use of Scientific Literature by Scientists,*" Rutgers University, November 21, 1956, 103 pp. Processed.
(12) Bureau of Applied Research, Columbia University, "*The Flow of Information among Scientists*, May 1958, 175 pp. plus appendices. Processed.
(13) "IBM System Disseminates Selected Data," *Electronic News*, July 17 (1961).
(14) Helen L. Brownson, personal communication.
(15) "Directory of Research-on-Research," College on Research and Development, Institute of Management Sciences, in preparation by the author.
(16) U. S. Air Research and Development Command, "*Vocabulary for Current ARDC Technical Efforts (CATE Program)*, February, 1960, 159 pp. Processed.
(17) David L. Marples, "The Decisions of Engineering Design," *IRE Transactions on Engineering Management*, Vol. *EM-8*, no. 2, June (1961).
(18) Jack Alan MacWatt, "Improving Scientific Communication," *Science, 134*, 313-316 August 4 (1961).
(19) "System of Future Gets Information to the Technical Men who Need It," *Petroleum Week*, March 1, 1957.

Measuring the Value of Information Services*

By JAMES HILLIER
RCA Laboratories, Princeton, New Jersey
Received September 29, 1961

Introduction.—Obviously, any discussion of information services by a member of the management of a laboratory will be influenced by the specific characteristics of that laboratory. It is, therefore, appropriate to describe the environment in which the following ideas on information services were generated.

RCA Laboratories probably is at the more complex end of the spectrum of industrial laboratories when considered from the point of view of information services. The 450 professional staff members to whom we would provide information services are working on more than 150 clearly identifiable projects. These projects involve at least 15-20 scientific and engineering disciplines. The research is primarily applied and exploratory but includes a substantial amount of basic work. It includes relatively little advanced development and no engineering.

In RCA Laboratories we have an outstanding but conventional technical library to satisfy the information needs of our technical staff. It is somewhat "outgoing" in that it does distribute to the technical staff copies of the tables of contents of all the current technical journals received. Recently we have been considering the possibility of establishing a more formal type of information service. Making a commitment to prepare a paper on the evaluation of information services is timely and has provided an additional incentive to give some serious thought to the subject. Unfortunately, the one result of that thought has been to raise the author's estimate of the complexity of the problem by several orders of magnitude and to uncover some significant dangers in an uninformed experimental approach. Incidentally, the conclusions reached here, largely on theoretical grounds, are very close to the conclusions reached in careful independent studies by individuals looking at the practical side of the problem in our laboratory.

In this paper an information service is considered to be an organization with a responsibility to take *positive* action toward satisfying the information needs of the technical staff it serves. It is distinguished from information retrieval systems (machine or manual) which are considered as responsive only to *requests* for specific information. A middle ground is possible, of course, in which the individuals providing information services to a specific research project are an integral part of the technical team working on the project.

A study of literature on the subject has been both confusing and disappointing. These seems to be a tendency to pick out a small part of the problem and make an intensive study of that part. This is a valid scientific approach, provided there is a good understanding of the whole problem and of the relationship the part bears to the whole. Unfortunately, the problem is too often vaguely described as "satisfying the user's need for information" without defining the need or appearing to

recognize the extreme ranges of variation that can exist in both the nature of the user and the nature of the information.

The first task, from the point of view of the management of a laboratory, is to establish the nature of the need. A laboratory engaged in true exploratory research can be considered as a "worst case" in establishing the nature of the need. For this reason it seems profitable to examine this aspect in some detail.

Information Needs for Exploratory Research.—Effective information flow in an industrial research laboratory doing exploratory research in a number of different scientific disciplines must satisfy several types of needs for information. The following list is one way of describing the major needs. The listing is in order of increasing importance to the success of the laboratory. The order probably bears no relationship to the cost of satisfying the individual needs nor to the volume of activity involved. Obviously, all of these functions are essential to the operation of such a laboratory. The information flow must:

1. Prevent excessive duplication of research, particularly of experiments destined to give negative results. This is a conventional information retrieval problem, since it can be assumed that a specific project or experiment is under consideration. It does, however, encounter problems of indexing and of vocabulary changes. The latter is a particular problem in new fields, such as certain aspects of electronics, where the precursor and relevant information is often found in several entirely different scientific disciplines. The "middle-ground" information service, that in which the information service is integral with the technical team, would avoid these problems.

2. Provide specific information needed by the technical staff. This category is intended to include the mass of technical and patent data and technical articles which usually can be specified accurately by the user. It involves routine information retrieval techniques except in those cases where there is uncertainty as to the exact source of the desired material.

3. Provide "catching-up" information for the individual who finds it necessary to become familiar with a new field. Obviously, since textbooks and review articles are primarily involved, the information retrieval problem is routine. Ideally, an information service group might increase efficiency if it had the skill to select appropriate items to match the individual's education experience, and the nature of the project to be undertaken.

4. Provide an efficient means for enabling the member of the technical staff to "keep current" in his project and his field. This requirement has never been handled effectively by conventional information retrieval techniques. For this reason, abstracting services have been established in different scientific disciplines. Unfortunately, the rapidly expanding volume of scientific publications and the increased range of scientific disciplines involved in many industrial research projects have greatly reduced the effectiveness of abstracting services as a tool for keeping current. This provides some justification for the development of laboratory-supported information services which concentrate on scanning and abstracting the literature of local interest and calling attention of the technical workers to publications which might be of interest to them. Determining the effective size and providing economic justification for this type of service are complex problems in which many factors must be balanced. Among the more obvious considerations are the size of the laboratory, the homogeneity of its research projects with regard to scientific disciplines, the average size of a project team, the rate of technical advance in the fields involved, the degree of competition at the research level, and the level of technical talent required in the service as compared to the level required to execute the research. It is discouraging to note how few of these can be measured quantitatively.

If we restrict ourselves to the need for keeping current, the basic justification for the existence of an information service group is the inability of the individual to survey the output in his technical field. By itself, this assumption can indicate the numerical strength of the information service group. This is simply the sum of the extra people needed to survey the literature in all of the fields of research active in the laboratory. If each member of the technical staff were working in a completely different field the information service group would obviously have to be larger than the technical staff. At the other extreme, when a whole laboratory is devoted to one narrow field the extra people needed to provide information to one member of the technical staff would serve the whole laboratory. It seems likely that intelligent estimates as to the necessary size of an information service group could be made in the case of a specific laboratory. Then management would have to exercise its judgment in balancing the cost of the service against its estimates of some of the other factors mentioned above. If keeping current were the only requirement, one could be optimistic about developing good judgment.

5. This leads to the fifth need, which is the crux of the problem in a research laboratory and particularly in one emphasizing exploratory research. This is the need for an information flow which will stimulate creative throught in a way which will maximize the probability of occurrence of creative ideas that are valuable to the company. Since the generation of creative ideas is the basic reason for the existence of any laboratory doing research of any level or kind, and since information is vital to creative thinking, this is the most important requirement of all. In fact it is not difficult to rationalize that those listed above are merely component parts of this requirement. While it is easy to make a simple and concise statement of this essential requirement, it is difficult (perhaps impossible) to define precisely how an information service can satisfy it. Obviously some benefits may be obtained by solving some of the component problems. There is the serious danger, however, that a lack of understanding of the most important requirement will allow the development of solutions to to the component problems which contain constraints on the creative process. For this reason it seems worthwhile to devote the remainder of this paper to a discussion of some of the basic concepts involved in the relationship between information flow and creativity.

Information Flow and Creativity.—When we attempt to establish some of the boundary conditions which must be taken into account in any attempt to understand the relationships between information flow and creativity and the relevant functions of research management, a useful guide is found in the similarity between human com-

munications systems and engineered communications systems as in the telephone, radio, television, *etc.*[1] Such a comparison develops several pertinent boundary conditions.

The most important characteristic of any communications channel is its capacity—the rate at which information can be passed through it. In practice, any communications system operates at or below its maximum capacity. In the case of the human communications system, the important point is that there is a limit to its capacity and that the limit is very small relative to the magnitude of the information flow in the environment of any individual. (Measurements of the channel capacity of human beings give values of a few tens of bits per second. However, for our purpose the actual value is unimportant.) The channel limit of the human system is all-inclusive. All of the information inputs and outputs to and from the individual must be included within his channel capacity.

It is limited capacity of the human system which prevents the scientist from absorbing all the literature that relates to his field of interest, even if he devoted all of his waking hours to the project. Considering that his limited channel must be shared with several other communications requirements of living and working, it is apparent that the individual must be satisfied with a restricted view of the literature and other significant information. It is, probably, the recognition of this situation that has led to the consideration of a preselecting organization interposed between the scientist and his conventional information retrieval system. Unfortunately, it is not possible to evaluate the effectiveness of such preselection until some of the other characteristics of the system have been analyzed.

Miller[2] and a group of his associates at the University of Michigan recently have made a study of what happens when a biological communications channel is overloaded. Their work uncovered a significant difference between the biological and engineered communications systems. They showed that when the capacity of a biological communications channel is exceeded by more than a modest amount, there is a sudden and complete collapse of the channel so that no information can pass through. Since evidence of such a channel collapse is not too common, we can reconcile the channel limitations of the individual human and the excessive amount of information available to him only by assuming that he exercises control, intelligent or subconsious, over the quantity of information he processes. One can go a step further and arrive at the conclusion that the control over the quantity of information processed is an individual matter developed by a lifetime of experience based on optimizing the probability of the individual attaining his personal desires, whetever they may be. This introduces the concept of quality of information, which is important relative to the attainment of personal desires. It appears that the individual combines the control of quality with the control of quantity by using a process of selection from among all the possible input and output communications channel in his environment.

In other words, the individual in a creative pursuit will have learned from experience the channels through which he has the highest probability of obtaining a useful piece of information and will tend to select those channels whenever a choice is involved.

To be effective in this framework, an information service will have to be able to demonstrate by actual performance that it provides a more profitable input channel to the creative individual. Moreover, it is the judgment of the individual rather than that of management that will prevail.

These thoughts begin to raise some questions as to the possible effectiveness of an information service.

Let us now turn to consideration of the process of creative action and to its impact on information flow to the creative individual. It is generally agreed that one of the essential conditions for the initiation of a creative action or idea is that the necessary information, or deductions therefrom, come together in time and in a single mind. It is also obvious by definition that these coincidences cannot be programmed—that is, neither the creative idea nor the information necessary to spark it can be specified in advance. This leads to the immediate conclusion that creative action, and therefore research, is probabilistic in nature. Thus, to the limitations and uncertainties imposed by a limited human communications system, we have added the further uncertainties of timing and combination in the creative individual's mind.

Management's responsibility in this situation becomes clear. In order to maintain or increase the productivity of the laboratory, it must find means of increasing the probability of creative action by the technical staff. Management can do this by increasing the individual probabilities involved. There are some techniques for accomplishing this, at least in part.[1] The specific question before us here, however, is that of the value of a scientific information service group placed either in series or in parallel with the scientist-information-retrieval link to preselect or filter the information flowing from the scientific literature. On the one hand, we have the need to find a way to display more of the ever-increasing mass of published scientific information to the creative individual in spite of the fixed limit on his communications channel. On the other hand, such control as there is over the probable occurrence of a creative action is in the hands of the creative individual and depends upon his lifetime of experience and upon his intimate personal goals. His success as a creative individual is a measure of his ability to optimize these probabilities without any *a priori* knowledge of the specific idea he will encounter or of the specific information needed to arrive at the idea.

Conclusion.—It becomes clear from this brief look at the major ingredients of the problem of providing information services to a research laboratory, that management's considerations must go beyond the simple provision of extra brains to perform a function that has grown to exceed the capability of the individual.

If we assume that the fixed channel capacity of the individual faced with an ever-increasing volume of literature leads to a reduction in the probability of creative action, we can conclude that there is a level of additional brain-power which will compensate for the reduced probability. However, we must recognize that the additional brain-power which is serving the creative individual must, in some way, operate in conjunction with the individual's experience-based optimization of his

information input. Otherwise, the service becomes a filter which reduces the probability of creative action and, furthermore, will be rejected by the individual.

We must also recognize that our assumption might be naive. The creative individual could very well have refined his selection process to such a degree that the increased literature available actually increases his probability of creative action, even though he never sees a large part of it. In this case, it is unlikely that any interposed system would improve the probability or, for that matter, be acceptable to the individual.

A further consideration is the balance between a centralized information service and a dispersed system, where the service is performed by individuals who are an integral part of the research teams. Here, too, the effects on the probability of creative action of the two approaches must be a most important factor in the evaluation.

The course of this discussion leads on to the conclusion that a qualitative understanding of the role of an information service activity can be developed on a rather general fundamental basis. However, it is clear that we do not yet have nearly enough detailed and quantitative data to enable management to make a reliable evaluation of an information service activity in a specific laboratory, particularly when that laboratory is a heterogeneous one doing exploratory research.

The discussion also suggests rather strongly that an empirical approach involves the danger of introducing unsuspected constraints which would reduce the probabilities of creative action.

The most obvious positive conclusion is that we must undertake further studies to develop a more complete understanding of the relationship between information flow and creativity.

* Presented before the Division of Chemical Literature, ACS National Meeting, Chicago, Ill., Sept. 6, 1961.

(1) J. Hillier, "A Theory of Communications in a Research Laboratory," *Research Management*, Vol. III, Number 4, pp. 255-270, Winter, 1960.
(2) J.G. Miller, *Am. J. Psychiatry*, **116**, 695 (1960).

Keeping Research in Contact with the Literature: Citation Indices and Beyond*

By JOHN W. TUKEY
Bell Telephone Laboratories, Murray Hill, N. J. and Princeton University, Princeton, N. J.
Received October 2, 1961

Varieties of Aim.—Those who write about information systems appear to believe that scientists go to libraries to get "information." This undoubtedly does occur, quite possibly much more frequently for scientists concerned with development. In many cases, however, especially in research, he goes to the library to *interact with the literature*. This is an active process, one in which the scientist's understood aims change steadily, both as a result of what he has found, and as a result of what he has accidentally noticed. The antithesis of the development man in search of specific information is the researcher who has come to browse.

It is not surprising that librarians and information systems proponents have concentrated upon the quest for specific information, for this is the library function which can be most completely transferred to another person, or even to a machine. Browsing, which can be least well transferred, is of key importance to keep research men alert and stimulated. Continual browsing can contribute much to the variety of the fields in which such a man can become expert on short notice. And with the exponential growth of science forcing more specialization on everyone, we should all be concerned with making it as easy as possible to be less specialized. Insofar as literature interaction for research scientists is concerned, we have to emphasize both easy browsing, and easy, rapid contact with the literature of a new sub-area. With these goals in mind, let us turn to some of the other aspects of the problem.

* Presented before the Division of Chemical Literature, ACS National Meeting, Chicago, Ill., Sept. 6, 1961.

Exponential Growth.—Today the most crucial fact about science is its continuing exponential growth (2, 7). Number of articles, number of journals, number of men and women, number of dollars, number of students, all are growing exponentially, doubling every few years, the exact doubling time varying somewhat from field to field and measure to measure. While exponential growth cannot continue indefinitely, and while the eventual slowdown will result in many new problems, it seems likely that research, the part of science most directed toward writing articles, will continue to grow exponentially for the longest time, as will the number of articles published.

Growth has already strained individual abstracting and indexing service to varying degrees. Many steps to help them are being considered, but the extent to which the services can surive is not yet determined. It is clear that we must not only try to help them, but must learn both how to supplement them and, perhaps, how to replace them. We need to keep as much of what we have had as we can afford, but we need to have new things as well. Both sheer volume and delayed indexing have made detailed searches using what we have today noticeably less effective. And the situation is going to get worse, either steadily or catastrophically. The needs are clear: (1) a way to dig into narrow fields; (2) a way to browse effectively; (3) a way to make exponential growth not a handicap, and; (4) a result that does not go out of data.

Non-obsolescence and the Literature Network.—If our means of focused access and easy browsing is to be slow to go out of date, it should depend upon relationship rather than upon classification. Times change, and the