

The Cost of Scientific Information*

JAMES W. RAMEY

Institute for Advancement of Medical Communication, University City Science Center,
Philadelphia, Pennsylvania 19104

Received May 10, 1965

I have been asked to discuss the cost of obtaining scientific information. After thinking this assignment over for a while, I began to appreciate how Alice must have felt playing croquet with the Queen. First of all, let me state my position clearly. I do not believe there exists, today a formula for putting a price tag on obtaining scientific information. I do believe, however, that we can define the parameters of the problem and pinpoint the issues to be resolved. I think we can develop crude, but revealing empirical means for determining the best trade-off between expenditures for certain types of information services and savings in scientific manpower—even though the basic relation between a scientist's use of information resulting from others' work and his own research productivity remains unknown. I shall suggest further a method, using available knowledge, by which I believe an organization can arrive at a gross estimate of its cost of obtaining information.

I think it is so difficult to place a dollar value on the process of obtaining information because basically we are unable to ascribe 100% of the cost of a document collection and the time the research scientist spends consulting it to the cost of obtaining "access" to information; also we often leave out of our estimates the factor of oral communication and its ancillary expenses. Finally, in scientific communication as in the use of document collection, we cannot assess accurately what percentage of the cost relates directly to securing information.

Several investigators have found that research scientists consistently place greater reliance on face-to-face communication rather than on communication through documents. The Case Institute study (1) did not specifically ask about reliance but did find that research scientists spend more time in face-to-face communication. Interaction is much richer, and speed and ability to get the needed information through give-and-take far outstrips extraction from documents. Menzel's study (2) of the information-seeking habits of university scientists indicated that sometimes information sought was unlikely to have been published or indexed. In such cases it would be essential to resort to oral communication. There is also the old problem cited by O'Connor (3) of inability to get a document out of a collection even when it has been identified because "it is on some reserve list, checked out, or buried in the basement."

Other imponderables inhibit our ability to relate directly our total use of the document collection to the cost of obtaining information. As Mueller (4) indicated in his study of a corporate information center, only a small percentage of retrievals may be of value—the others being either false drops or references to other documents. Also, if in the process of looking for information, one realizes that he has incorrectly defined the problem, incorrectly phrased the question, or started at the wrong place, it is doubtful that this human error can be considered a cost of obtaining information. If, on searching for material on question A, one stumbles onto an answer to question H, a new lead for a student assistant, a shortcut for a friend at the next bench, or a new recipe for rum punch, can these unexpected bonuses be properly offset against the cost of finding information? I think not.

Even if we agree to lump together, as legitimate costs of obtaining information, all expenses of maintaining and staffing the information center (which in industry amount to 1.5 to 2.5% of the total research budget according to a survey by Herner, *et al.* (5)), plus the total time research scientists spend consulting the collection and other written sources, it would still be necessary to assess oral means of getting information. (The Case Institute study found chemists spending twice as much time in oral scientific communication as in reading.) Here we have the same basic difficulty in assessing costs. Whether we talk about formal oral communication methods (such as attending meetings) or informal oral means (such as telephone and face-to-face conversations), participation in professional activities, use of consultants, or allowing staff members to be consultants, the same question can always be asked: how much of the cost should actually be charged to the cost of obtaining information? While most of these activities also serve other purposes they are nevertheless legitimate information-getting activities. It might even be instructive to ask which dollar is most productively spent: for journals, long distance calls, another information specialist, or perhaps the addition of a coordinate indexing system to the document collection. This question is not raised often enough and when it is raised, it is side-stepped handily, particularly in research proposal budget justifications. There are probably few principal investigators who have not undergone the unnerving experience of having visitors from one of the Federal funding agencies cast a jaundiced eye on already pitifully small budget requests for books, subscriptions, information specialists, library supplies, telephone, postage, and travel and proceed to question these requests as though they were shameless attempts to rob the public till. Clearly, some portion

* Presented before the Division of Chemical Literature, Symposium on Work and Time Studies in Technical Information, 149th National Meeting of the American Chemical Society, Detroit, Mich., April 1965.

must be allowed, but establishing the amount is a real poser. Any formula one might concoct for this purpose would necessarily be based on pure speculation.

If one is willing to exclude these philosophical imponderables, one could arbitrarily add up the total cost of the information center, all subscriptions, telephone bills, travel, consultants' costs, cost of allowing members of research staffs to be consultants, postage, formal and *ad hoc* meetings, and participation in professional activities and conclude that the resultant dollar figure is the out-of-pocket cost of obtaining information. This would be a sizeable chunk of the total research budget. Add to this the Case Institute figure of 32% of the salary cost of the researchers, while they are involved in these information-seeking activities, and the resulting figure could run as high as 40% of the total research budget. After one has this figure, what does it mean? Lacking knowledge of how expenditures for these activities are related to the output of the research effort, one cannot say that any figure is too little or too much. Attempting to separate the cost in dollars and time of the information-gathering activities of scientists, both formal and informal, from the overall cost of their creative effort is a sticky problem—one which again makes our cost figures meaningless.

While I do not doubt that sooner or later someone will manage to conduct a rigorous research study with proper control of the myriad and complex variables involved and suitable control groups, which will determine the effect of various types of information service on creativity, much more research will be needed before we can feel secure about assigning cost figures to these functions.

In the meantime, even if it is not possible to assign precisely the cost of each method of obtaining information, it may be possible to show that one means is more expensive than another. We would thus be establishing a gross value hierarchy, even though precise cost measurements were lacking. By constructing a table that combines the average cost of the chemists' time spent on reading (9.5%), using the Case study figures, with total information center costs (postage and subscriptions, his general discussion time (7.5%) with attending formal and *ad hoc* conferences and travel, and his oral nondiscussion communication time (11%) with total telephone costs, consulting costs, and use of personal files), it would be possible, in very gross terms, to compare the cost of these three means of obtaining information—but unless these figures can be related to a unit of information, such as cost per question answered, these comparisons will still be meaningless.

Even after developing such crude cost figures, we are still faced with a value problem that goes beyond dollar

expenditure. The researchers themselves value certain information-gathering means far out of proportion to their relative cost. Nevertheless, it may be that even they (the creative group we hope to assist) may not be the final arbiters of this question. A recent MIT (6) study of the utilization of information sources during the preparation of research proposals indicates that some of the factors involved in obtaining information may be very heavily weighted, at least in the minds of those who approve research contracts, without regard to their relative cost to the contractor. For example, technically competent full-time staff people are apparently valued much more than outside consultants. If it could be shown that this is generally the case among those who decide which proposals to approve, the question, "what price, scientific information?" becomes even more meaningless.

It would seem more meaningful to me to ask such operationally definable questions as what is the unit cost of an information service per *answer*, rather than per search? Furthermore, what is the unit cost per answer *delivered* vs. unit cost per answer *accepted*? To carry this line of inquiry even further, we could ask the unit cost per question satisfied? A similar question can be posed with regard to ancillary services, for example, routing. Routing can be evaluated in terms of unit cost per new idea or previously unknown document discovered, that is, in terms of novelty. Hopefully as a result of chipping away at the problem through getting answers to such questions as these, we will one day be able to answer the question, "what is our cost of obtaining information per unit of productivity?" This is really the \$64 question.

LITERATURE CITED

- (1) Anon., "An Operations Research Study of the Scientific Activities of Chemists," Case Institute of Technology, Operations Research Group, Cleveland, Ohio, 1958.
- (2) Menzel, H., "Planned and Unplanned Scientific Communication," Proceedings of the International Conference on Scientific Information, National Academy of Sciences-National Research Council, Washington, D. C., 1959, Vol. 1, pp. 199.
- (3) O'Connor, J., *IRE Trans. Eng. Writing Speech*, **5**, 75 (1962).
- (4) Mueller, M. W., "Time, Cost and Value Factors in Information Retrieval," Lockheed Aircraft Corp., Burbank, Calif., 1959.
- (5) Herner, S., *et al.*, *Ind. Eng. Chem.*, **51**, 56a (1959).
- (6) Allen, T. J., "The Utilization of Information Sources During R & D Proposal Preparation: Research Program on the Organization of R & D," Alfred P. Sloan School of Management, Cambridge, Mass., 1965.