

Second Thoughts on Scientific Information

BY ALVIN M. WEINBERG

THE WEINBERG of the so-called Weinberg Report would appear to take his life in hand when he agrees to speak before a group of librarians since I have been the librarians' *bête noire* ever since the report was issued. There is a rumor that a librarian in California ceremonially burned *Science, Government, and Information*.¹

I should point out that, though the report criticized librarians, it criticized scientists much more severely. The librarian was depicted as a responsible citizen valiantly trying to cope as best he knows how with the flood of scientific information; the scientist was depicted as shirking an essential part of his responsibilities in the handling of information. I therefore propose that we bury the hatchet and get on with the job at hand. The information crisis is so severe that, even with *fullest* cooperation between the information and scientific communities, we can hardly expect to resolve it.

I shall concern myself here with what I consider to be one of the main messages of *Science, Government, and Information*: namely that science, in response to the information crisis, is undergoing a hierarchal social reorganization, and that this social reorganization will impose a corresponding hierarchal organization on our scientific information system. The central elements in this organization I see as the information center and one of its chief customers, the theoretical scientist. In particular my remarks will be concerned with the relation between the information centers and the

Mr. Weinberg is Director of the Oak Ridge National Laboratory. This paper was prepared for presentation before the Association of College and Research Libraries and the American Library Association, St. Louis, Missouri, June 30, 1964.

librarian. In elaborating on this matter I shall not claim to speak as an expert on information but simply as a scientific administrator who is as puzzled as anyone about how to keep science from crumbling in the face of its information crisis.

THE CRISIS IN SCIENCE

Our panel concluded that science—as a consistent view of nature—is in danger of fragmenting into mutually contradictory pieces as its explosive growth overtakes and clogs its communication system. An example of such mutual inconsistency was pointed out last year by M. King Hubbert, then president of the Geological Society of America.² Apparently hydrologists had for many years used an equation to describe the flow of incompressible fluids that left out an important term; hydrodynamicists of course knew the right equation but, because hydrologists did not speak much with hydrodynamicists, this point in theoretical hydrology contradicted the laws of hydrodynamics.

The crisis, as our panel saw it, is sufficiently acute to require drastic action by all communities concerned with science and scientific communication. The information people had already been alerted:

¹ U. S. President's Science Advisory Committee, *Science, Government, and Information* (Washington: Govt. Print. Off., 1963), 52p.

² "Are We Retrogressing in Science?" *Science*, CXXXIX (March 8, 1963), 884-90.

we therefore addressed a large part of our report to the technical community. The report was viewed, as much as anything, as a means of alerting the scientists themselves to the responsibilities they must assume if we are to have much hope of coping with the information crisis. As we said in the first few sentences of the report, "Transfer of information is an inseparable part of research and development. All those concerned with research and development—individual scientists and engineers, industrial and academic research establishments, technical societies, government agencies—must accept responsibility for the transfer of information in the same degree and spirit that they accept responsibility for research and development itself."

In this respect I think the response to our report has been disappointing. Members of the information community—librarians in particular—are all familiar with *Science, Government, and Information*; the government agencies have been in rather a tizzy about information since the report appeared. On the other hand, the technical people—those who most needed to be alerted—seem to have been much less influenced by the document. I would be surprised if as much as five per cent of the technical community knows about the report, whereas I have yet to meet a technical librarian who has not read at least a review of it. Perhaps the librarians could bring the report to the attention of their clientele; I know the panelists would be grateful if you did this.

Some of the specific recommendations and observations in the report had to do with the substance of the information problem; others had to do with organizations that the government has set up to handle information. I suppose that the most striking observation about the information problem was one introduced by Professor Eugene P. Wigner, recent Nobel Laureate in Physics and a member of our panel. Professor Wigner views science as undergoing a social reorganiza-

tion in response to its growing problem of communication. He sees the scientific community layering into several hierarchies. At the first level are the bench scientists, each of whom works in a rather narrow field, and each of whom communicates with other closely related bench scientists. The results of each group of bench scientists are, in Wigner's view, kept under surveillance by the next group of scientists—the group leaders or bosses. These group leaders communicate with each other, and in this respect maintain contact between different groups of bench scientists. In principle, the hierarchy could be extended with groups of group leaders themselves being kept under surveillance by supergroup leaders who again would communicate with each other at a higher level of abstraction than do the group leaders. This proposed hierarchal structure for science corresponds to a separation into different levels of abstraction that are reminiscent of Alfred Korzybski's "Structural Differential." The traditional working scientists are at the bottom rung—each one knows almost everything about almost nothing; as one progresses toward the top of the pyramid, the subject matter becomes more abstract until one finally reaches the philosopher at the top who knows almost nothing about almost everything.

In some branches of science the theoretical scientist, who works at a higher level of abstraction than do the bench scientists, has manifestly taken on this job of higher-order surveillance. I had a good opportunity to see this illustrated recently during a meeting on the chemical basis of mutagenesis, sponsored by our biologists in Gatlinburg, Tennessee. Most of the conference participants were experimental biochemists; they knew each other and each other's work. Among the group was the distinguished French quantum chemist, B. Pullman, whose role at the meeting was crucial. It was he, the theorist, who on the one hand seemed to know the experimental data better than did any of

the other participants; and on the other hand, it was he who tried in every instance to fit each piece of data into a broader framework. Moreover, Professor Pullman maintains contact with the community of quantum organic chemists, as well as with the community of biochemists. He is a good example of Wigner's group leader who interprets and correlates the results obtained by the individuals in his group (the bench biochemists), and maintains relations with other group leaders in neighboring sciences, *i.e.*, in quantum chemistry.

The theoretical physicist does this sort of job for physics; to a lesser extent, the theoretical chemist does the job for chemistry; and to an even lesser extent, the theoretical biologist (such as Professor Pullman) does it for biology. Now the main point of all this is that the theoretical physicist plays a central role in resolving the information crisis in physics. His theories, on the one hand, deepen our understanding of the experimental phenomena; but on the other hand, as Edward Teller has stressed, they provide a simplified framework upon which one can hang broad reaches of experimental data. For example, there are now more than eighty elementary particles—mesons, strange particles, and strange particle resonances—and the number is growing relentlessly. For anyone who is not actively working in elementary particle physics, and even for those who are, just to keep track of the particles—to know the difference between a Ξ^0 and a Λ^- , or to remember the isotopic spin of the K^- meson—is a formidable task. But recent theoretical studies have uncovered striking regularities among these particles—the most recent being the Ne'eman-Gell-Mann eightfold way—which not only deepen our insight into the particles but also enable us to keep them straight in our minds. In this sense the theorist, in correlating and providing a conceptual framework for isolated experimental facts, achieves much the same end as does the information

specialist; he, in effect, compacts the literature much more efficiently than can any abstracting service or even review article.

There are many other examples of how the theorist compacts the literature. Perhaps the best example from chemistry is the Mendeleev periodic system and Bohr's *auf-bau prinzip* which explains the periodic system in terms of quantum theory. The regularities embodied in Mendeleev's chart immediately make vast reaches of chemical data available to the average chemist. Before Mendeleev the chemist had to remember that the hydroxides of Li, Na, K, Rb, Cs are strong bases; after Mendeleev he needed only to remember that the bases of the Group I metals are strong, and that Li, Na, K, Rb, and Cs belong to Group I, a fact that he would want to remember for other reasons. In this way Mendeleev compacted the literature tremendously for many generations of chemists. Bohr's insight represents a further compaction of the data: Bohr gives the key to understanding how to construct the Mendeleev table. Moreover, Bohr, plus the ideas of quantum chemistry, enables one to reconstruct the chemical properties of elements—to predict that the hydroxides of Group I metals are strong bases without having to remember this specific fact. The great power of quantum theory in compacting the chemical literature explains why some of the newer high school chemistry courses are based on the chemical bond approach. In the Wigner hierarchal scheme, Mendeleev would be one level above the bench chemists; Bohr would be at a still higher level.

A similar over-all compaction of the literature in biology was made possible by evolution and, more recently, by the Watson-Crick model of DNA. In each case one can, to some extent, predict regularities—and therefore recall facts—from a knowledge of the underlying theory.

The theorist as compactor of the literature is imperfect. He is most successful in the simple sciences, like physics, less successful in the complicated sciences like

chemistry or biology, and hardly at all successful in the supercomplicated behavioral and social sciences. This is only a way of saying that theoretical science has become a well-defined technique only in the simple sciences, and its usefulness, either as a means for achieving deeper insights, or as a way for compacting the literature, decreases as the subject matter becomes more complicated.

Yet the theorist in his role as compactor of the literature is so useful that we are naturally impelled to try to invent devices or approaches that can play, in the sciences too complex to be compacted by theory, the role that the theorist as literature compactor plays in the simple sciences. In other words, we ask, can the process of compacting the literature, by identification of easily remembered regularities in complicated situations, be systematized?

What I am describing here is really the inductive method in science: the correlation of many seemingly disparate facts and the identification of regularity in a sea of diversity. Perhaps by examining some of the great inductive insights of past science, we may be able to learn how to encourage and systematize induction.

Possibly the greatest piece of scientific induction was Darwin's formulation of the theory of evolution and of natural selection. Darwin's method was induction at its very best—a painstaking, deliberate collection of vast amounts of data, a persistent and thoroughgoing study of the data, and a brilliant identification of regularities that constitute the theory of evolution. Once the theory of evolution had been stated, large stretches of biology, taxonomy, and paleontology could be systematized: an impossibly complicated field acquired regular features and, in this sense, became easier to comprehend.

A second major example of successful scientific induction is the formulation of the shell model of nuclei by Jensen and Mayer, for which these two physicists shared a Nobel Prize last year. Many of

you know that our picture of the nucleus was dramatically changed by the observation that clusters of neutrons or protons with the "magic" numbers 2, 50, 82, and 126 are very stable. Nuclei containing this many neutrons or protons play for nuclear structure the same role that the rare gases with 2, 10, 18, 36, etc., electrons play for atomic structure. Though such regularities were suspected very early by Elsasser, the full implication of these early speculations had to wait until large amounts of data on properties of nuclei had been amassed and correlated. And in fact, collection of the growing mass of nuclear data became an important occupation for many nuclear physicists; had the data not been collected, it is hard to see how Jensen and Mayer could have formulated the shell structure of the nuclei. Now that a sort of Mendeleev table for nuclei is available, the literature of nuclear science in effect is compacted. For example, one remembers that the capture of neutrons by lead is improbable because lead has a magic number of protons, 82, and *all* magic nuclei have a small tendency to capture neutrons.

The successful inductions all share the same pattern: the data is amassed; it is systematized; someone worries very hard and very long about it and, with luck, discovers the regularities. In the old days of Darwin, or even Mendeleev, the amount of data that was to be handled was sufficiently small so that it was possible for the same person, or a small group of persons, to be bench scientist, amasser of data, correlater of data, and inventor of new syntheses. But the situation is very different now, especially in the complicated sciences. The data which scientists add to our existing store of data, the new facts and observations, are a huge flood. A division of labor between those who create or discover the facts and those who sift, absorb, and correlate the facts seems to be inevitable. This inevitable division of labor corresponds very nearly to the social layering of science envisaged

by Wigner. The scientists who collect and sift the facts would in our modern technology constitute an information center. In the simple sciences like physics, which already have a well-defined theoretical structure, the output of information centers is fed to the theoretical scientists who try to make the generalizations that encompass all the data. In the complicated sciences where the role of the theoretical scientist is less well developed, the output of information center is utilized as best the science can to find elementary generalizations and simple syntheses.

In a sense, then, the information center is the key element in the systematization of the process of induction in modern, complicated sciences. As we see it, the center, to be successful, must be complete: it must have in its system *all* the relevant data. Moreover, it must select and interpret; it must therefore be manned by scientists who not only handle the documents but also glean the scientific gems from the documents and try to make sense out of them. The people who run the centers must be recognized scientific leaders in their fields of science—in fact, people like K. Way at the Oak Ridge Nuclear Data Center or Y. S. Touloukian at the Purdue Thermophysical Properties Research Center or Raymond Pepinsky at Florida Atlantic University Crystallography Center, are scientists of high order in their own right who, in the process of examining the data, find the new correlations and in this respect create new science.

RELATION BETWEEN THE INFORMATION CENTER AND THE LIBRARIAN

The specialized information center in our view is therefore not a technical library: it is more nearly a technical institute since in its ideal form it creates new science, in the way that Darwin and Mendeleev created new science through their wonderful inductions based on vast reaches of data. But the specialized center obviously must depend on the librar-

ian in the most central way if it is to operate successfully. First are the obvious and traditional library functions—particularly retrieval, storing, and cataloging. Every information center must have on its staff librarians who are expert in these fields.

But the demands of the information center are rather special and go somewhat beyond the traditional demands made on the librarian. The information center is expected to achieve about 99.5 per cent retrieval of the literature in the special field in which it operates. The 80 per cent coverage tolerated by research scientists is not condoned in an information center for otherwise it will lose the confidence of its public. The Nuclear Data Project, which searches the literature of physics and chemistry most carefully, recently missed a report on a half-life measurement of Ra²²⁸ in reports of the College of Medicine of Utah State University. This oversight was pointed out by a leading physicist who condoned it to a certain extent but who clearly would have thought better of the project had it not occurred.

How can 99.5 per cent retrieval be achieved, even in a limited field, without demanding endlessly tedious, expensive, and overlapping literature searches by highly trained scientists? Experiments are being started on prepublication indexing of articles by authors, referees, or other specialists. Librarians could contribute their wisdom and experience to the development and support of the needed high-retrieval systems.

The other great problem of the information center is to make its products—state-of-art reviews, compilations, analyses, etc.—known to those who would use it. None of the present publication outlets seem suitable. For example, publishers are reluctant to advertise such specialized products; journals do not want long tables; as a result, compilations are appearing mostly as laboratory reports. Titles and designations like those shown in the following table are not uncommon.

No doubt any librarian in a big library can find such reports, but for the scientist in a small college in Pakistan the difficulties are great. The experience of librarians could help much in finding some good general solution to the problem of marketing and distributing compilations.

RECENT COMPILATIONS OF NUCLEAR DATA

Comprehensive list of nuclides with atomic mass, half-life and specific activity	AHSB-44
Cross sections for fast neutron reactions	Texas Nuclear Corp.
Radionuclides arranged by gamma ray energy	Nucleonics Data Sheets
Coefficients of internal conversion of gamma radiation, Part II: L-shell	USSR-Press; Moscow, Leningrad; issued in USA as 58ICCL1
Radiations from radioactive atoms in frequent use	USAEC-February 1959
Effective cross section values for well-moderated thermal reactor spectra	CRRP-960 EANDC-4 TNCC-30 AECL-1101

Another problem in this area in which librarians' help is sorely needed is the problem of steering an information-seeker to the appropriate compilation, if it exists, rather than to the primary literature. If someone asked you for information on the boiling point of water you would direct him to a handbook or compilation. If someone asked you for information on the nuclear magnetic moments of isotopes of silver, would you be able to guide him properly?

You might perhaps spend a while looking in various indexes and abstract services under Silver, or Isotopes, or Nuclear Magnetic Moments, or Moments, Nuclear Magnetic. You would find some sixteen articles (that is, after winnowing out the extraneous material), most of them in the *Physical Review*, and in perhaps three or four hours your information-seeker could go through these papers and find the magnetic moments he needs.

But this would be an awful waste of your time and his because if you directed him instead to a compilation by Ingmar Lindgren he would find in a few seconds the nuclear magnetic moments of seven isotopes of silver, the methods of determination, and the references.

Whenever the desired information has been compiled it is better to refer the information-seeker *first* to the compilation, then, if necessary, to the primary literature.

Where librarians' help is needed is in devising systems for quickly finding what has been compiled and where to get a copy.

If the desired information does not exist in compiled or analyzed form but is included in the subject field of some information center it is often better to refer the information seeker to that information center rather than to attack the primary literature. The new National Referral Center for Science and Technology, which is being developed at the Library of Congress can be a great help in connecting potential users with information centers.

THE LOCATION OF THE INFORMATION CENTER

Since in our view the information center is a technical institute primarily rather than a technical library, it seemed natural to us to propose that the information centers be housed where science is actively going on—that is, in large laboratories or in universities. We at Oak Ridge have taken the advice of our panel very seriously indeed, and we have established six

different full-fledged information centers at ORNL—in Research Materials, Nuclear Safety, Radiation Shielding, Isotopes, Nuclear Data, and Charged Particle Cross Sections, as well as several smaller ones.

Each of our information centers is distinctive, reflecting as it does the personality of its leader and the character and state of the field of science and technology it covers. Thus I have already alluded to the Nuclear Data Project, headed by K. Way. This center has operated for the past sixteen years, until last year in Washington under the aegis of the National Research Council. During this period, the group operated as an independent and somewhat isolated center. However, Dr. Way decided that the group would gain strength from closer contact with active research in nuclear physics—that scientists working at the Nuclear Data Center would bring a finer discrimination and judgment to their management of the nuclear data for being at a place where nuclear experiments are going on. So to speak, the two levels of Wigner's hierarchy, represented by the working experimental scientist and by the compiler and sifter of data, reinforce each other by being in close contact.

The data collected at the Nuclear Data Project is well defined, and it fits neatly into the broader, conceptual framework of theoretical nuclear physics. At the other extreme of the spectrum is our Nuclear Safety Information Center, headed by W. B. Cottrell. Here the subject matter—safety of nuclear reactors—is much more amorphous, much less quantitative, much less readily placed in a conceptual framework than is nuclear data. The compaction of the literature achieved by this center is achieved more by the state-of-the-art review than by the brilliant generalizations or profoundly based theory. Yet many of the modes of operation of the two centers resemble each other. The leader of the NSIC, Mr. Cottrell, is one of our country's foremost experts in nu-

clear safety, just as Dr. Way is in nuclear structure. He and others who work in the center also carry out an extensive experimental program in the field of nuclear safety; they add to our knowledge of nuclear safety both by correlating and compacting the work of others and by seeking out new knowledge themselves.

THE GOVERNMENT AND THE INFORMATION CENTER

I should like now to turn to the impact that the information center is having on the government's handling of scientific information. Many of you may be aware that the report, *Science, Government, and Information*, was prompted to some extent by the criticism that Senator Humphrey's Subcommittee had been leveling for years at the way the executive branch had been handling scientific information. The government has reacted with considerable vigor to many of the suggestions in our report—information seems to have become the concern of the management of government agencies where previously it had been something that only librarians and information specialists were concerned with. The idea of the specialized information center has caught on, and one sees a tendency for the agencies to encourage their proliferation. The Department of Defense for example has set out to define just what is a specialized information center more closely than was done in our report; it has taken an inventory of the centers it has been sponsoring; and it seems prepared to establish new centers.

I believe that the information center may eventually help resolve one of the organizational problems that plagues the government's support of scientific information. Some of you may recall that in our report we urged that government agencies with partial responsibility for research in certain areas of science should pool or coordinate their information activities in the field, and that one agency be designated the delegated agent for information activities in the field. As the

delegated agent, the agency would accept responsibility for seeing that the proper things were done in information pertaining to that field: where relevant journals needed support the agency would provide such support; where information centers were needed the agency would provide them, etc. In this respect information in each field of science would become the responsibility of a single agency which would godfather that field; the agency would serve as a little Office of Science Information Service for a field of science in much the same way that NSF's OSIS godfathers science information as a whole.

The delegated agent idea has encountered opposition, and for an obvious reason. Agencies which were responsible for certain segments of research were understandably reluctant to give up responsibility for information in this field to another agency—there was always the danger that the next step after relinquishing control of information about a field would be relinquishing control of research in the field—and this would mean a cut in budget, something no agency can contemplate with equanimity. The delegated agent idea therefore has been all but killed—the only exception being atmospheric pollution where the United States Public Health Service has been designated delegated agent.

I believe, though, that the information center may achieve at least a part of the rationalization of information-handling envisaged in the concept of the delegated agent, yet avoid the political problems. The advantage of a single rather than multiple management of a body of information is that it is much easier on the user. If, for example, abstracting of the work on lasers is done independently by the seven or so agencies that support work on lasers, then the poor worker in the field has to look in seven places, not one, to keep up with what is being done. Now suppose the abstracting were done by an information center concerned with lasers. The information center would keep under

control work on lasers regardless of where the work originated or regardless of who paid for it—in this respect, the information center imposes uniformity on the flow of information on lasers that captures much of the essence of the delegated-agent idea. So to speak, by consolidating the management of information at a subagency level—*i.e.*, at the specialized information center—and supporting the information center with money from many agencies, one achieves the simplicity which the delegated agent seeks to achieve, yet avoids the political difficulties.

THE FUTURE OF THE INFORMATION CENTER

I shall close my remarks by voicing whatever concerns I have about the trend toward information centers, a trend which *Science, Government, and Information* has undoubtedly accelerated. On paper the information center seems like a good idea; it appears to me to be a neat practical embodiment of Professor Wigner's social layering of science; that the information centers sprang up spontaneously suggests that they in fact serve a clearly felt need.

And yet there are pitfalls which must be avoided if the centers are to play the central role we envisaged. Information centers are expensive: one is dedicating a perceptible fraction of one's scientific manpower to a second-order scientific activity in the hope that out of such second-order activity will come better science. Now the most difficult thing is to assess just how useful are the information centers—do they indeed fulfill the goals which we have envisaged? would science be poorer for not having them? would they spring up spontaneously again if they were closed? These questions are not unique to evaluation of information centers. They are questions that arise whenever one tries to decide how much information is needed, in fact, whenever one supports a service which, before the

service was established, one had lived, or better existed, without. Information is in many ways like advertising—one knows that some advertising is needed; the problem always is to figure out how effective your advertising is, and how much you can reduce your advertising budget without suffering loss of business. So it is with the information center—one must guard against mudpie-making, and one must be assured that there is a real need for the products; although, as in the case of the citation index, the needs may not always be recognized by the customer. Possibly canvassing the center's customers regularly would be useful; or perhaps measuring the volume of the complaints that arise when an information center reduces its business; or subjecting a number of

older information centers to a critical appraisal. These are vague ideas, and I hope others will occur to this community of librarians.

But I do not want to leave you with the wrong impression. So far I see no better mechanism than the information center to sift the information flow in the newly organized Wignerian scientific hierarchy. I hope I have made clear to you that, as the information center grows, so will the library that supports the center—and that, in our coming scientific information system, this example of cooperation between librarian and scientist in operation of the information centers will serve as a pattern for the much broader cooperation between scientist and librarian that our ever-growing science will demand. ■■

ALA Representatives to Academic Ceremonies

CARSON W. BENNETT, Heidelberg College, Tiffin, Ohio, represented ALA at the inauguration of William Travers Jerome, III as president of Bowling Green State University on September 15; Richard H. Shoemaker, Rutgers University, at an academic convocation commemorating the silver jubilee of Caldwell College on September 19; Bernard Kreissman, City College of New York, at the inauguration of Ralph Gordon Hoxie as chancellor of Long Island University on October 9; Brother Brendan, St. Michael's College, at the dedication of St. Johns College in Santa Fe, New Mexico, on October 10; Charles F. Gosnell, New York University, at the inauguration of Alan Simpson as president of Vassar College on October 16; Margaret Thomas, Randolph-Macon Woman's College, Lynchburg, Va., at the inauguration of Marion Carey Brewer as president of the college on October 17; Helen Stockert, West Virginia Wesleyan College, Buckhannon, at the inauguration of Gordon E. Hermanson as president of Davis and Elkins College on October 21; Sister M. Regis, Immaculate Heart College, Los Angeles, at the academic convocation in connection with the golden jubilee of Jesuit higher education, on October 22; Ralph H. Parker, University of Missouri, at the dedication of James Madison Wood Quadrangle, Stephens College, on October 22; Lewis M. Ice, University of Bridgeport, at the installation of William C. Innes as president of Fairfield University on October 24; Richard H. Logsdon, Columbia University, New York City, at the inauguration of Albert H. Bowker as chancellor of City University of New York on November 5; Lyman S. Tyler, Brigham Young University, Provo, Utah, at the inauguration of James C. Fletcher as president of the University of Utah on November 5 and 6; Robert D. Harvey, Southwest Missouri State College, Springfield, at the inauguration of Ernest S. Brandenburg as president of Drury College on November 6; Jean Henderson McFarland, Vassar College, at the dedication of the U.S. Military Academy library on November 13. ■■