Experimental Data as a Source for the History of Science

CLARK A. ELLIOTT

The Memorabilia was full of ancient words, ancient formulae, ancient reflections of meaning, detached from minds that had died long ago, when a different sort of society had passed into oblivion. There was little of it that could still be understood. Certain papers seemed as meaningless as a Breviary would seem to a shaman of the nomad tribes.... The Memorabilia could not, of itself, generate a revival of ancient science or high civilization, however, for cultures were begotten by the tribes of Man, not by musty tomes; but the books could help, Dom Pedro hoped—the books could point out directions and offer hints to a newly evolving science.[1]

WALTER MILLER'S NOVEL of a reviving civilization in a post-atomic world has a special message for the archivist—the central role of records in the transmission of scientific knowledge. Archivists, when called upon to evaluate the records of science, may feel at times like the monks in this story. These records become the "Memorabilia" which we blindly accept and revere but never really understand, hoping someone will come along and unlock the secret. As often, perhaps, we violate the basic motivations of the Order of St. Leibowitz and reject the records of science entirely just because we do *not* understand them.

The historian of science shares with other historians an interest in correspondence, diaries, administrative records, published and unpublished works, and the like. Historians of different disciplines, too, have their own unique records. Literary historians, for example, often have a real need for successive manuscript drafts of a creative work; economic historians may use accounting ledgers and

The author is Assistant Curator at the Harvard University Archives. This paper is an expansion of one dealing with archival potentials of ADP media that he read at the annual meeting of the Society of American Archivists in St. Louis, September 28, 1973.

1 Walter Miller, *A Canticle for Leibowitz. A Novel* (London: Weidenfeld & Nicolson, 1959), p. 144.

journals. The history of science has "laboratory notebooks," a generic term meaning a record of experiments. This record may include descriptions of those experiments, test data, computations, comments and evaluations of experimental results, and much else, and it can be in separate notebooks or may even be on random scraps of paper. These notes are the primary records for the documentation of the history of scientific development. Upon the test data and laboratory notebooks rest, successively, correspondence, progress and research reports, and published accounts.[2] In a less formalized structure, there may be personal journals or scientific diaries, memos, and the like. Notebooks, however, are at the base of experimental science documentation, and other records are summaries at different levels or in different contexts.

No one doubts the importance of science in the twentieth century, and yet the history of science in this century has been little explored, either by professional historians of science or by historians of contemporary society.[3] Partly, perhaps largely, the explanation for this lack of investigation lies in the fact that historians (including historians of science) are generally ill-prepared to deal with the technical aspects of science.[4] Among the varieties of the history of science are what Thomas Kuhn categorized broadly as intellectual history and the history of the socioeconomic role of science.[5] It is in writing the history of science as a subdivision of intellectual history that historians must consider the use of laboratory notebooks and other data sources. Kuhn implied this usage when he wrote: "When scientific ideas are discussed without reference to the concrete technical problems against which they were forged, what results is a decidedly misleading notion of the way in which scientific theories develop and impinge on their extrascientific environment."[6]

If adequate sources are going to be available to historians to write the history of twentieth century science, archivists must begin to act now—largely without the optimum technical background that it would be desirable to have. In a second article, Thomas Kuhn discussed the historic role of quantification in science and pointed out

2 Alice Yanosko Chamis, "Managing Tech Records. Variety of Records Calls for a Variety of Techniques and Tools," *Information and Records Management* 4 (April/May 1970): 18–19. This article describes the situation in the library of the B. F. Goodrich Research Center. In a university setting, recordkeeping is frequently much less formal.

3 Margaret Gowing, "Science and the Modern Historian," *Times Literary Supplement* 69 (7 May 1970), p. 515.

4 Thomas S. Kuhn, "Relations Between History and History of Science," *Daedalus* 100 (Spring 1971): 276–78.

5 Ibid.; see also p. 283.

6 Ibid., p. 279.

that measurement usually follows theory.[7] The important role of measurement is to confirm or question a scientific theory, and in this regard scientists often will go to great lengths to make their data fit the theory.[8] Much of the work on the methodology or philosophy of science comes from physics. Quantification has a somewhat different role in other sciences. In chemistry, for example, quantification is largely useful as a means of classification and distinction and, as one writer put it, "The aim is not, as with physics, the enunciation and verification of general laws or the precise description of generalized processes."[9] Quantification (counting) is used in a similar though more simplified way in taxonomic biology.[10] Even in other aspects of biology, however, measurement "derives from, feeds into, sharpens and clarifies, and discriminates between alternative qualitative descriptions and models; it cannot generate them."[11]

The important generalization for the archivist, it seems, is that the data of measurement by itself and out of its theoretical context is of no interest to the historian. The higher levels of scientific documentation—correspondence, scientific diaries, progress reports, publications—must take priority insofar as they shed light on the mental processes of the scientist and the interaction of theory and data. Many laboratory notebooks are particularly useful in that they will serve both of these functions—to preserve the original data and to show how the scientist dealt with it. Notebooks or other compilations of raw data and calculations would be considerably less useful without supporting theoretical documents.

Conflicting opinion exists about the adequacy of published documents as sources of scientific data and as documentation for the history of science in its more technical aspects. T. R. Schellenberg believed that in general the published record adequately met researchers' needs.[12] One article, based on experience with the records of the British Atomic Energy Authority, suggested that only exceptionally interesting notebooks and working papers need be kept "since most scientific work is written up in reports."[13] Kendall

7 Thomas S. Kuhn, "The Function of Measurement in Modern Physical Science," in Harry Woolf, ed., *Quantification: A History of the Meaning of Measurement in the Natural and Social Sciences* (Indianapolis and New York: The Bobbs-Merrill Co., 1961), p. 37.

8 Ibid., pp. 41, 50.

9 Henry Guerlac, "Quantification in Chemistry," ibid., p. 67.

10 Ibid.

11 R. W. Gerard, Quantification in Biology," ibid., p. 204.

12 Meyer H. Fishbein, "A Viewpoint on Appraisal of National Records," *American Archivist* 33 (April 1970): 179.

13 "The Records of Science and Technology, with Thoughts on Their Disposal," *Archives* 8 (April 1967): 29.

Birr concluded that for the history of ideas, the published record would usually be adequate.[14] Presumably the history of ideas, if it can be broadened to mean intellectual history, is the area that would be most concerned with the technical development of science, and yet Birr does not prescribe notebooks as a necessary source for this kind of history. Maynard Brichford, predicting that research in the history of science will come to depend increasingly on published sources, wrote that "Test and experimental data should be destroyed when the information they contain is condensed in published reports or statistical summaries."[15] And the American Institute of Physics' Project on the History of Recent Physics, taking a somewhat different view, concluded that "It is one of the unavoidable shortcomings of modern scientific publication that details of instruments and calculations are increasingly being omitted in the effort to make publications as brief as possible. As this policy develops, the historic and the scientific value of laboratory notebooks is enhanced considerably, even when they are largely numerical."[16]

Obviously, there are differences of opinion regarding the usefulness of notebooks and data sources compared to published sources for the history of scientific development. Few would debate the need for at least a summary of the data and a description of how it was derived, but if less is now available in print there seems little choice but to keep at least some of the data in manuscript form. Even if publications do contain adequate summary data for purposes of communicating the results of experimental work, there is still the question of whether this is adequate for the level of understanding which an historian of science might seek. He will need to comprehend fully the nature of important experimental work—how it was carried out, what the results were, and how successive experiments approximated the theoretical expectations. The nature of historical inquiry is that the historian cannot merely assume a progression from uncertainty to success. He knows that the history of science is characterized by trials and errors and by periods of relative calm and of revolution in scientific thought, and he wants to know these byways as well as the major progressive steps. The problem ultimately is to know the relative importance of a person or a project in

14 Kendall Birr, "'What Shall We Save?' An Historian's View," *Isis* 53 (March 1962): 7.

15 Maynard J. Brichford, *Scientific and Technological Documentation: Archival Evaluation and Processing of University Records Relating to Science and Technology* (Urbana-Champaign: University of Illinois, 1969), pp. 6, 17.

16 Project on the History of Recent Physics in the United States, Publication R-152, *Scientific Instruments and Apparatus: Sources for the Fuller Documentation of the History of Physics* (New York: American Institute of Physics, June 1963), p. 7.

the history of scientific thought and the likelihood of their scientific records being used as historical documents.

While the scientific work of every scientist will not be studied in detail, and therefore the notebooks of every scientist need not be kept, there are other more conventional records, such as correspondence, which are of immense value in studying the communication of scientific ideas, the workings of the scientific community, and the relationship of the scientist to society.[17] The outcome might be to collect notebooks, correspondence, and other records for the important scientists and to begin the collections at the correspondence level for minor scientists.

Albert Michelson, Nobel Prize winner in physics, is a man of great scientific importance. D. T. McAllister has described Michelson as a man who wrote for those scientists who shared his own points of reference. Michelson did not feel the need to fill his scientific articles with details of experimental data, calculations, or apparatus and techniques, and yet now the Michelson Museum wants to collect any and all bits of paper relating to Michelson's scientific career, including apparatus, raw data, notes, and calculations.[18] In this instance, it appears that these records of detail, of Michelson's day-to-day activity and his thoughts in relation to his continuing scientific work are of immense value to historians, whereas they were of little value to Michelson's contemporaries in science. This poses the very real question of the relative value of notebooks and data records for the historian and the fellow scientist.

Archivist of the United States Wayne C. Grover thought in 1962 that records of science were of value not only to the history of science or to biographies of scientists, but also to the advance of science.[19] There is also the possibility that an archives may be asked to store the data records of a scientist who hopes that he or someone else will someday be able to exploit them fully for scientific work.[20] Nathan Reingold did *not* think that scientists are likely to use the data records in archives, of other scientists. He stated emphatically, "Scientists simply do not work that way."[21] It does seem very

17 "The Records of Science and Technology," *Archives:* 28.

18 D. Theodore McAllister, "Collecting Archives for the History of Science," *American Archivist* 32 (October 1969): 327, 333.

19 Wayne C. Grover, "The Role of the Archivist in the Preservation of Scientific Records," *Isis* 53 (March 1962): 58.

20 J. Frank Cook, "The Archivist: Link Between Scientist and Historian," *American Archivist* 34 (October 1971): 377; and A. Hunter Dupree, "What Manuscripts the Historian Wants Saved," *Isis* 53 (March 1962): 66, and his "Comments" at the Conference on Science Manuscripts in this same special issue of *Isis*: 94.

21 Nathan Reingold, "The National Archives and the History of Science in America," *Isis* 46 (1955): 23.

unlikely that any scientist would come to a conventional archives to consult, in conjunction with his own current work, the data of an earlier scientist. Except in a formal sense, he probably does not even consult the very early published documentation. At times there may be special bodies of data which have continuing scientific value but not historical value. Libraries should be encouraged to accommodate these stores of data rather than sending them to an archives. In the long run, conventional archives can probably best serve the advance of science by supporting the documentation of the history of science rather than scientific research.

It is well to keep in mind that much of the discussion in this paper has referred directly or indirectly to laboratory science. Reingold has pointed out the special considerations which apply to evaluating documents which record observational data in their natural setting—for example, geographical, zoological, botanical, meteorological, and astronomical observations. The observations can be visual and descriptive or by scientific instruments, but they usually cannot be reproduced since they come directly from nature. These "natural history" records[22] frequently do not fit a pattern of observation that is carried on to confirm a theory. The records are gathered and employed in descriptive science or are studied statistically in the hope of inductively observing patterns in nature. Their value will frequently be quite beyond the theoretical; indeed, without them the scientific information which they display may not be known. They are much more the necessary ingredient in the advance of such a level of science than are the data gathered or generated to support a theory.

Computer applications in the sciences may necessitate the need to question fundamentally and perhaps to reevaluate all that has been said above about the data records of science. There is an important historical difference in the use of the computer for scientific purposes and for commercial or business purposes. Generally speaking, in the sciences the computer is used more for computation, while in commercial applications storage and access are the more important and useful features of the computer.[23] Another aspect of this differentiation appears to be of particular importance to the archivist. As one writer put it: "Relatively few mathematical or scientific tasks require, by their nature, to be repeated in their entirety. Considerable preparatory work may be necessary before a mathematical data processing task can be performed on a computer. Once the job is

22 Ibid., p. 24.

23 Keith London, *Introduction to Computers* (London: Faber and Faber, Ltd., 1968), pp. 64–66.

run on the computer successfully it is not usually repeated. In commercial applications, however, there are many repetitive processes."[24] The implication is that computer language records in the sciences have only a fleeting or temporary value and are merely a means to an end. The same might be said for other scientific data records, and yet in their proper theoretical context they have great value for the historian of scientific thought.

One author has pointed out that "computers have become part of science because they permit altogether new ways of obtaining and using information."[25] It is now possible to handle large amounts of data on weather, geophysics, biology, and ecology which could not be handled before.[26] There are also computer applications of a different order which allow the human mind broader dimensions in dealing with natural or experimental phenomena. One example is the use of an electron beam to draw pictures of physical phenomena which in turn can be photographed and studied.[27] The data and programming documentation for important projects of this nature would be of great interest to historians both of science and of the methodology of investigation. The use of computers in studying signals from space, or the study of sub-atomic particles on earth, allows the compilation and manipulation of enormous amounts of data.[28]

It is possible for a scientist to conduct experiments in which the scientific data is read directly by a computer from the scientific apparatus and subsequently analyzed or calculated by computer.[29] In such a case there never is a paper record until the printout, and the printout may include or involve only selected data. The equivalent of the laboratory notebook in this situation is the magnetic tape alone. Just as with data in conventional laboratory notebooks, the matter of theoretical significance, the importance of the project or the scientist, the degree of documentation of experimental design, and the degree to which the data is reported in published reports and journal articles must be weighed in deciding whether to retain it in an archives for the history of science. In addition, the archivist

24 Ibid., p. 243.

25 W. O. Baker, "Computers as Information-Processing Machines in Modern Science," *Daedalus* 99 (Fall 1970): 481.

26 Ibid., p. 486.

27 Ibid., p. 488.

28 Ibid., p. 490.

29 John Caffrey and Charles J. Mosmann, *Computers on Campus: A Report to the President on Their Use and Management* (Washington, D.C.: American Council on Education, 1967), pp. 33, 71.

must preserve the complete computer programming and coding record in human-readable format.

Computer-language records, just as laboratory notebooks and data records of other kinds, must be considered for use by both historians and scientists. It will be some time before historians get to the point where they will use such sources. But when they do, these records should be waiting for them. It may be that larger stores of data can be preserved than in the past just because of the computer's greater compactness and versatility. It is ultimately preferable, where both paper record and machine-readable record exist, to retain the machine-readable record because it can be viewed in different ways, a virtue that a printout does not have. Historically speaking, it would also be important to know in which form—that is, in what configuration—the scientist actually looked at his data in matching it with his theoretical expectations. In this instance it might be necessary to keep his printouts or at least the ability to recreate them. The historian would want also any manuscript notebooks or logs in which the scientist registered the successive experiments or "runs" and his comments on them in light of the theoretical context of his experiment.

Beyond the needs of historians, archivists might want to look more broadly at the possible interest of other scientists in the data. Given a certain amount of uniformity in data collection and storage, it is probably more likely that machine-language data will be of greater interest to other scientists than have a miscellany of very personalized laboratory notebooks or other manual systems of data registration.

The extent to which much of this discussion is strictly hypothetical becomes apparent when one looks at the documentation cited in the history-of-science literature. In surveying the last twenty-five years of *Isis,* the Journal of the History of Science Society, for articles dealing with nineteenth and twentieth-century American science, I had the impression that published sources, both primary and secondary, were extensively cited, manuscripts less frequently. Citations of manuscripts were usually references to letters, and sometimes to minutes, reports, diaries, and some miscellaneous lists and other documents—but almost never to laboratory notebooks or data sources. The almost total neglect of such sources by American historians contributing to *Isis* is surprising. It would be a truly worthwhile project to analyze statistically the documentation cited in a broader range of works on the history of American science to see what the patterns of use have been. I checked also several union catalogs of manuscript collections in the sciences[30] and came away sensing that manuscript repositories have not yet fully faced up to the special problems of documentation for the history of science. Certainly correspondence files and other such generalized sources are fundamental, and they perhaps show a wider swath of scientific life than do a laboratory notebook or a data file. But the life of experimental and observational science resides as much in the laboratory and field as it does in the study or at the typewriter. To emphasize one record over the other distorts the way that science and scientists truly work, and this is a disservice to a fundamental force in contemporary society.

30 American Institute of Physics, Center for History and Philosophy of Physics, *A Selection of Manuscript Collections at American Repositories,* ed. by Joan Nelson Warnow (New York: A.I.P., 1969); also, Edwin T. Layton, Jr., ed. *A Regional Union Catalogue of Manuscripts Relating to the History of Science and Technology Located in Indiana, Michigan, and Ohio,* Publication No. 1, Program in the History of Science and Technology (Cleveland: Case Western Reserve University, [1971]), the "Repository" listings; and scientific and technical records reported in the *American Archivist* "News Notes" during the years 1967–71.