

## Chapter 28. Data Gathering and Methodology in Writing: Fact Production and Use

My initial study of environmental information, discussed in the previous chapter, continued my interest in the production, representation, and circulation of facts through texts and how inscribed facts are consequential for individuals and societies. Growing up anticipating a career in the sciences, I had come to appreciate evidence and data. Even as my interests turned to the humanities, I maintained commitment to evidence, including texts as evidence in literary studies. My commitment to facts became more self-conscious as a graduate student when my dissertation advisor, J.V. Cunningham, emphasized that a few simple facts about text production, finances of the publishing industry, or historical and cultural conditions could rein in fanciful, speculative literary interpretation.

Nonetheless, as I became more aware of the complexities of fact production and particularly as I engaged with science studies, I needed to grapple with the realization that facts were humanly made and were social constructions. They existed as formulations within human communications. How could facts and even data be humanly produced and yet still provide reliable information about the world? This problem articulated at least as early as David Hume in the eighteenth century was reignited in the theory and science wars of the 1970s and '80s. I soon recognized, however, rather than this paradox being a philosophic puzzle or the dividing line of an ideological battle, it identified sites that could be examined empirically to see how data and evidence were produced and warranted in various different social locations. The social construction of facts offered sites for empirical investigation of how epistemology was managed practically within knowledge producing collectives.

As I turned my attention to academic and disciplinary writing, I saw how the content of disciplinary texts depended on the methods of data collection and analysis viewed as credible in each field. At end of '70s when I started work on "What Written Knowledge Does," while representation of the world was considered one of components of writing (recognized by its place in the rhetorical triangle), I could find no rhetorical study of the role of disciplinary differences of evidence and data, let alone methods, as scientific writing was then largely considered outside the realm of rhetoric. As I examined texts from biosciences, sociology, and literary studies, however, I noticed differences in the kinds of evidence used, where the evidence came from, and how the evidence was made relevant to the argument.

When I came to revise subsequent editions of *The Informed Writer*, I tried to capture some of these disciplinary differences in gathering and representing

evidence (see Chapter 18). The sources of the data cut across disparate disciplines in surprising ways. Working with historical artifacts puts geology, evolutionary biology, paleontology, history, archeology, and literary studies into the same pot. Observing contemporary events connects journalism, sociology, anthropology, astronomy, descriptive plant and animal biology, and ecology. Creating experimental events to produce specific kinds of data is of course common in medicine, physics and chemistry, but also in branches of psychology and sociology. Largely theoretic fields, such as philosophy and mathematics, foreground ideas and claims from other texts. Most fields, nonetheless, engage in some abstract reasoning, when evidentiary bases for claims retreat to the background. This approach to where evidence could be found and how it was collected got me thinking about method and epistemology.

## The Representation and Analysis of Data

The questions of data, where they come from, and how they are evaluated kept popping up as I studied different instances of scientific writing. In examining Arthur Holly Compton's notebooks (see Chapter 19), I saw how his writing was motivated by the desire to present a new kind of data made possible by a newly invented instrument, the bubble chamber. This method of producing data raised many questions since the visible, recordable data were only photographs of trails condensed by moving particles. From these trails one could infer the kinds of particles, their energies, and their angle of deflection from the initial collision. After calculations based on these assumptions, the photos of these trails were taken as evidence of quantum theory. Some of the photographs were judged by Compton to be defective or to raise other problems, so he excluded them from the final results. His notebooks were careful in recording, evaluating, and interpreting the angles—which then provided the basis of the article (along with a description and rationale of the methods and the theory which was corroborated by this new form of evidence).

Data and method came up again in studying *The Physical Review*, as the kind of data collected and the set of related methods producing them in optical spectroscopy became my constant against which I could identify other kinds of changes in the discourse. Data and method came up again in the history of *The Philosophical Transactions* as evidence became presented more precisely and quantitatively while the conditions producing the data became more specified. Over time, measurement tools improved and experiments became designed more intentionally, based on expanding methodological reasoning. Method was at issue in the way Newton represented the process of his observing and gathering optical phenomena in his various writings. I found Joseph Priestley paying detailed attention to all the previous evidence of electrical phenomena going back millennia, as well as to the evolving machinery that produced and recorded electrical charges. He also recommended that every investigator specify exactly what they did and why,

even when experiments failed. Then more recently in the social sciences, I found that the American Psychological Association *Publication Manual* in each revision increased specification in the methods sections in order to restrict what would count as professionally credible data.

The question of credible and relevant data was the central obstacle in knowledge traveling readily between the two apparently close specialties of toxicology and ecotoxicology. Because toxicology was a laboratory experimental medical specialty and ecotoxicology was a field-based statistical study, differences in sources of evidence and methods used to gather and analyze the data were sources of resistance to the credibility and usefulness of knowledge transferring across the fields. These differences of methods reflected underlying ideologies and epistemologies. Ultimately, however, the need for the findings of the other field to deal with problems within each led to greater acceptance across specialties, accompanied by a broadening of the epistemologies and ideologies of both.

In studying how pharmacological findings entered into court proceedings in the US, I found the barriers even higher and the procedures for crossing boundaries even more systematic and complex, due to the differences of the structure and purposes of legal and scientific institutions, creating radically different approaches to evidence. Understanding and representing the differences in the institutions was a necessary part of explaining the logic of the procedures surrounding the translation of scientific knowledge into the evidentiary and judgment procedures of the courts. Later, in studying how scientific knowledge of climate change entered Congressional hearings, I found the process controlled not only by the nature of the institutions, but the complexity of interests and motives of legislators and the strategies used by majority parties to support their political agendas. Another earlier study of educational assessment and policy concerned the oversized attention given to learning outcomes that were easily collected and counted, in contrast to the difficulty of gathering data that recognized and supported the accomplishments of progressive educational practices (Bazerman, 2003e).

More personally I experienced the importance of methods, data, and the difference of collection practices as I attempted to publish interdisciplinary research in venues of different fields. I quickly discovered that journals from different fields would not consider manuscripts credible or even relevant if they did not follow those fields' expectations and methods; I needed to obtain field-specific data with field-specific methods positioned within the relevant theories and reasoning of the field. No matter how well I carried out the theoretical work, if I didn't argue from the right kind of evidence, the work would not be seen as worth attention. So although writing was crucial to the objects of study of many fields, it had to be addressed in each field's terms.

In my own field of writing studies, after a first round of research methods books appeared around 1990, little emerged to update or replace them. This led to Paul Prior and myself editing a collection on text analysis methods (Bazerman & Prior, 2004i). In order to assure the consistency needed in a textbook, we

provided a chapter template and expectations to guide the authors, each of whom were expert in the methods presented in their chapters. Each of the chapters focused on one set of methods, and together they indicated how every act of writing brought together simultaneously multiple dimensions. Although humanities traditions of text analysis had a history of eclectic and implicit interpretive practices, we thought it useful to bring their long-standing analytical methods to greater explicitness and to put them into the context of methods from the social sciences, so that researchers could consider which analytic tools best fit their inquiries. Doing this work made me more sensitive to the analytic tools of my colleagues and expanded my own analytic repertoire. There have been similar benefits from my role as a regular seminar leader at the Dartmouth Summer Seminar in Composition Research from its founding in 2011 until now.

## How Undergraduates Learn Methods of Data Production and Analysis

These various experiences heightened my curiosity about the processes by which writers came to learn and use these methods, and what the consequences were for their understanding of the phenomena they were looking at. I remained haunted by a diagram in Latour and Woolgar's *Laboratory Life* (1979, p. 46). This diagram suggested how through continuing processes of inscription, the laboratory was in essence a factory for turning mice (along with other inputs like energy, laboratory equipment, chemicals, and office supplies) into scientific articles to be submitted to journals. Latour and Woolgar saw this process of repeated inscriptions as a process of progressive forgetting, obscuring the material lives of the mice, and by extension the world studied. But I also was interested in what was being remembered, focused on, recorded, calculated and reasoned about—the non-fictionality as I discussed in a review I published when *Laboratory Life* first appeared (Bazerman, 1980a). Having something new, revealing, and evidentially persuasive to share with one's disciplinary colleagues seemed to me to be at the heart of academic writing. And learning the methods to produce and think about such evidentiary data was central to disciplinary training and effective disciplinary writing.

To study the learning of methods, I first considered looking at working scientists, but their mature practices and assumptions were deeply tacit, embedded in their view of doing good science, so it would be hard to excavate. Vygotsky had suggested that the best place to study a process was when it was just being formed and was at the foreground of behavior and attention rather than later when it was sedimented and taken for granted (Vygotsky, 1978). I next thought of studying how graduate students learned methods, but I decided to look earlier at undergraduate students when they were first being explicitly introduced to the practices of disciplinary inquiry. At that moment misunderstandings, missteps,

corrections, and guidance from their mentors might be most visible. While much undergraduate writing I suspected (accurately, it turned out) would be dominated by textbook presentations of already established knowledge, perhaps supplemented by cookbook lab experiences, yet, as students specialized in their majors, they might be asked to form their own inquiries and gather fresh data. As I began to speak with colleagues about their disciplinary curricula, I found that the senior thesis was often the first place that students produced their own data. Locating useful sites for study turned out to be as much a matter of prepared serendipity as earlier location of and access to specialized archives. That is, I needed to know the kinds of situations I was looking for, and then be able to spot them when they came into view. In this case I had to develop ways of talking to people to see what opportunities emerged.

The first place I could gain access to a senior project was in the mechanical engineering program at a neighboring university. I had gotten to know the coauthor, the mechanical engineering professor who supervised the project, through his participation in one of my seminars. The year-long senior team project proposed, designed, prototyped, and tested in both a lab and the field a low-cost foot prosthesis to be produced at a rural clinic in Honduras. From the four major reports the students prepared, along with related reflective documents and interviews, we could see how they identified the kinds of data they needed to produce for each report and how they went about producing and collecting the data—drawing on the library, site visits, theoretical calculations, active fabrication processes, and lab and field testing. As the students were designing and then fabricating a device which was then tested, they were bringing new objects into the world based on prior fact-finding; these new objects were then the sources of further data. All of these forms of data provided the substance of their reports, and selected data and findings from earlier reports would get carried forward into later reports as authoritative knowledge and assumptions for the next level of work and production. This showed us how data get built into continuing action, innovation, and production. The largest writing challenge of this project was to create an unfolding narrative of the building of knowledge across the year studied, showing how the exigencies of each report required specific kinds of data collection, which then became sedimented into later stages of the project, guiding the later decisions, data-gathering, and reporting (Bazerman & Self, 2017f).

In the next study I wanted to focus more centrally on students' perceptions of the methodological challenges they faced across the year as they developed their projects. Through speaking with people aware of curricular developments across my campus, I learned which departments were currently engaged in self-study of their undergraduate programs and would appreciate more information about how their students learned to write in their fields. After meeting with representatives of several departments, I found the Political Science department to be particularly welcoming and interested in finding out more about writing throughout their program and in their senior honors seminar. With the permission of the

chair and members of the department, I pursued the research on a double track, with a study (based on syllabi and faculty interviews) of writing in all the undergraduate political science courses and a more focused interview and text study of self-selected students in the senior honors seminar. The first study was for an institutional report to meet the department's needs, but it provided useful background for the more focused study. In contrast to the engineering study where the reports were the center of the story, the written projects of the senior political science honors students served only as evidence of the consequences of the work students were doing to meet these perceived challenges, which they described during periodic interviews. It turned out the students with richer practical methodological experience and abstract methodological understanding were able to look more deeply into the material they were studying and were able to write more effective and subtle analyses in their final reports. Their final work was exactly constrained and focused by the methodological knowledge that they had. The resulting article again followed the unfolding narrative, but this time with a more phenomenological focus, developing each student's perspective on what they were doing, what challenges they faced, how they were trying to address them, and what they had accomplished at each point (Bazerman, 2019c).

The last study in this series emerged when I was doing an independent study with upper division undergraduates on the issue of data and evidence. All three undergraduates, by chance, were linguistics majors and we started talking about the linguistics courses where they actually worked with a lot of data. They mentioned one instructor who in several different introductory-level courses gave numerous assignments asking students to analyze and interpret data provided by the instructor. This seemed to me a good opportunity to study the impact of the working with data on students' understanding and perception of language. After confirming the teacher's use of such material with colleagues in the Linguistics Department, I found out more about these assignments from the instructor and identified a course he was willing to have studied. Valentina Fahler, a graduate student I was working with at the time, also had a strong background in linguistics, so I asked her to collaborate on this project. The instructor and two of the four Teaching Assistants in the course volunteered to be interviewed—the instructor about the design and intentions of the course and the assignments, and the TAs on what they were observing about the students' struggles and development across the term. The main data, however, came from the students who volunteered to fill out regular questionnaires and provide their assignments. We found, however, student responses on the questionnaire were constrained by the format of questions asked and did not indicate striking changes across the year. Further, since the assignments throughout the year drew on different kinds of materials and asked different kinds of questions as part of the instructor's well-designed developmental sequence, it was not easy to make simple comparisons across the term. Further, as with any group of students, they brought different knowledge and experiences to the class, so it was not easy to aggregate or compare their work

quantitatively. However, in going over the data with Valentina multiple times as we did the coding, we formed individualized pictures of the students. When we combined all the data sources including their assignment submissions, we found that we could get a consistent picture of development of each of them, filtered through their individual situations. We could also see some similarities within clusters of students. We went through multiple iterations of descriptions of the development of each student, cross-checking our perceptions for negotiated inter-reader reliability. Eventually we were able to offer a group phenomenological interpretation relying on both individual and small cluster portraits. As we revised towards publication, we had to tighten these accounts and become more selective in the detail to make the portraits readable and the larger patterns more visible (Fahler & Bazerman, 2019f).

These last three studies all presented challenges of choosing and integrating multiple data sources for contextual, phenomenological, and performance evidence. The analyses and presentations needed differing balances and relations among these forms of data, according to the nature of each inquiry. This inquiry into learning data practices is far from complete, not even getting yet to an interim synthesis and articulation. Nonetheless, I hope these studies indicate that data collection and analysis methods are directly relevant for writing studies and writing education, and that quality of many kinds of writing is in part related to the quality of finding and inscribing primary materials to report on and analyze. Students can write better if they have good stuff to write about, and they can draw on more persuasive evidence if they know how to produce and analyze it according to disciplinary standards.