Inscribing the World into Knowledge:
Data and Evidence in Disciplinary
Academic Writing

Charles Bazerman*

Abstract
The production of data, their proper analysis, and appropriate use as evidence are at the heart of academic writing; further, students’ enculturation into disciplinary-appropriate practices of evidence use is central to the development of disciplinary competence. Writing Studies research, however, has much still to learn about the process of inscription of data (that is, how data is produced and recorded so as to be available for analysis and calculation), the way the data then becomes evidence deployed in academic writing, and the form the evidence takes in the written products of different disciplines.

This chapter examines the challenges faced by three university students majoring in political science as they work on their senior honors theses. Overall the student interviews suggest that the prior training and experience in the gathering and manipulation of data affected numerous parts of the thesis writing process. The prior experience has an effect on the final thesis, including the formation of the research question, the flexibility, and variety of data gathering methods conceived and deployed, the precision of implementation, the kind and nature of discovery made in the project, and the understanding of the complexity of phenomena investigated. Further, in this instance, the prior learning of methods and development of methodological sophistication was not primarily the result of an organized curriculum but was based on idiosyncratic individual experiences. The idiosyncrasy of experience in this

* Distinguished Professor of Education, Gevirtz Graduate School of Education, University of California Santa Barbara.
E-mail: bazerman@education.ucsb.edu
study heightened the differences among the three students, thereby making more visible the relationship between previously learned methodological skills and writing practices, processes, and results.

**Keywords:** data and evidence, academic writing, higher education, disciplinary methods

**Resumen**

La producción de datos, su adecuado análisis, y su uso apropiado como evidencia está en el centro de la escritura académica; además, la enculturación de los estudiantes en prácticas disciplinarias apropiadas de uso de evidencia es fundamental para el desarrollo de la competencia disciplinaria. Sin embargo, la investigación en Estudios de Escritura tiene mucho que aprender sobre el proceso de inscripción de datos (es decir, cómo se producen y registran los datos para estar disponibles para su análisis y cálculo), la forma en que los datos se convierten en evidencia desplegada en la escritura académica, y luego la forma que la evidencia recoge en los productos escritos de diferentes disciplinas.

Este capítulo examina los desafíos que enfrentan tres estudiantes universitarios que se especializan en ciencias políticas mientras trabajan en sus tesis honoríficas. En general, las entrevistas estudiantiles sugieren que la formación previa y la experiencia en la recolección y manipulación de datos afectaron a numerosas partes del proceso de redacción de la tesis, dando lugar a diferentes resultados en la tesis final, con consecuencias para la formación de la pregunta de investigación, la flexibilidad y variedad de datos, los métodos de recolección concebidos y desplegados, la precisión de la implementación, el tipo y naturaleza del descubrimiento realizado en el proyecto y la comprensión de la complejidad de los fenómenos investigados. Además, en este caso particular, el aprendizaje previo de los métodos y el desarrollo de la sofisticación metodológica no fue principalmente el resultado de un plan de estudios organizado, sino que se basó en experiencias individuales idiosincráticas. La idiosincrasia de la experiencia en este estudio aumentó las diferencias entre los tres estudiantes, haciendo más visible la relación entre la experiencia metodológica y las prácticas, procesos y resultados de la escritura.

**Palabras clave:** datos y evidencia, escritura académica, educación superior, métodos disciplinarios

Our knowledge of disciplinary writing practices has advanced greatly since early analyses of writing in different subject domains (Bazerman, 1981). We
know more about the form of academic genres (Swales, 1990) and their rhetorical function (Miller, 1984) and their role in activity systems (Russell, 1997; Bazerman, 1994). We have learned about how these genres change in different historical contexts (Bazerman, 1988; Gross, Harmon & Reidy, 2002; Atkinson, 1998). We have also learned about the complex uses of intertextuality and how they are tied to genres, engagement in disciplinary knowledge and strategic display (Devitt, 1991; Berkenkotter Huckin & Ackerman, 1991; Bazerman, 1993) as well as about disciplinary and academic identities (Castelló & Donahue, 2012) and academic agency (Schryer et al., 2002).

An area that has been more elusive has been data—their identification, collection, selection, display, analysis and use as evidence in disciplinary and professional arguments, and how these aspects of data production, manipulation, and use are differentiated in different domains and activity systems. Writing Studies as well as Science Studies have paid some attention to the alternatives for alphabetic, graphic, or mathematical representation, (for example, Gross & Harmon, 2014; Coopman et al., 2014; Hentschell, 2014; Kimball, 2013; and Hutto, 2008), but have not looked much into how the data are actually are produced or disciplinary differences in production, as the focus has been on representation. Some might say that such matters belong to specialized disciplinary practices and are rightfully categorized as issues of method and methodology—and thus are outside the domain of writing specialists and investigators. Indeed, these matters of methods have been extensively discussed within disciplines, and these discussions have been precisely the site where fields have developed their distinctive ways of interacting in the world to inscribe data reportable in articles and used as credible evidence for claims, theories, accounts, and projected designs. The reasoning and evaluation that warrant or discount various methods are central to the reliability and credibility of empirical endeavors and form central parts of the disciplinary work. Further, training neophytes into the disciplinary methods of inquiry, data production, and data-grounded analysis and argument is a central part of forming disciplinary and professional specialists. So, there are good reasons for writing researchers to be cautious in poaching in these well-guarded disciplinary preserves.

Still, all writing is about stuff—sometimes imaginary stuff or conceptual stuff or obscured stuff—but always about stuff, evoking some meanings, thoughts, representations, calculations, or feelings of the reader. The quality of the writing is often centrally evaluated on what it tells us about the world and how compellingly and persuasively. In disciplinary terms, these
evaluations and the credibility of its meanings almost always rely on how the stuff of the world is captured or represented in the data. How those data representations are collected from the world, how they are presented in the text and then analyzed, and how that data then become evidence in reasoning are essential to warranting arguments in empirical disciplines.

Students approaching assignments typically look for good topics—that is, topics they know about, that allow access to information appropriate to the task and discipline, and that they know how to disciplinary reasoning. If they write about interesting stuff represented incredible and warrantable ways, they are more likely to make discoveries and to come to fresh new thoughts. They know if they have the right stuff to write about and can get the data or evidence about it, they are likely to write more compellingly for the instructor and be evaluated more highly. As well they are more likely to be pleased with their writing and feel they have learned and grown in their thinking.

Subject matter teachers, as well, are likely to be less concerned with the correctness of language than whether disciplinary knowledge and concepts are understood and applied accurately and whether students collect, analyze and use disciplinary-appropriate data and evidence in ways appropriate to the discipline. Other aspects of writing may help sharpen, clarify or make the arguments more understandable, but the ultimate criterion is whether students can see and represent the world through the lenses of disciplinary method and theory.

So if we are to help students improve their writing and engage more effectively in their disciplinary and professional tasks, we do need to engage with how they choose, gather, and represent the stuff they talk about and how they use those representations to support and elaborate their ideas. That is, we need to engage with methodology and epistemology. This study examines how the extent of theoretical and practical knowledge students have about methods and methodology affect how they go about completing a major senior research project that requires the gathering and use of data, and how the resulting paper shows the effect of that methodological know-how. The three students in the honors program in a single major examined here share little structured training in methods but rely on widely different experiences before or outside the major, affecting how they go about their senior project and the resulting theses.

Of course, epistemology and methodology are fraught issues, with long histories of discussions that are at the heart of scientific and scholarly identity, dating back to at least to Francis Bacon’s *Novum Organum* published in 1620.
The heart of the problem is that the stuff of the world can never be put directly into writing. The problem is well captured by Jonathan Swift in 1735 in the third journey of *Gulliver’s Travels* when Gulliver encounters the Academy of Laputa, where some savants propose to improve knowledge by a Scheme for entirely abolishing all Words … since Words are only Names for Things, it would be more convenient for all Men to carry about them, such Things as were necessary to express the particular Business they are to discourse on…many of the most Learned and Wise adhere to the New Scheme of expressing themselves by Things, which hath only this Inconvenience attending it, that if a Man’s Business be very great, and of various kinds, he must be obliged in Proportion to carry a greater bundle of Things upon his Back, unless he can afford one or two strong Servants to attend him. (Swift, 1906, p. 170)

The real issue is not the strength of the savants’ backs or those of their servants, but whether they can do away with representation. A published paper cannot contain actual rocks or birds or human cultures, but only representations of them. These representations may include charts or photos or videos or in the future even virtual reenactments—but still they are only representations. And at some point, in the text, even these representations need to be discussed in the symbolic terms of natural language perhaps supplemented by mathematical formulations. As noted above, the literature in both writing studies and science studies have focused on the manipulation and interpretation of symbols within texts, both to improve the intelligibility and rhetorical effectiveness of texts and to highlight the production of knowledge as a symbolic meaning within the social and psychological worlds mediated by texts. This literature, however, tends to adopt a perspective polar to that of scientists—that the representations are entirely of human construction and do not bear any particularly compelling relation to the material worlds they claim to present.

So, the puzzle remains that objects and phenomena of the world need to be transformed into symbols, with all the limitations, ambiguities, and manipulations that symbols are heir to, as well as all the reductions if they are to be formed as human knowledge. As Alfred Korzybski (1933) notably said, the map is not the territory—but texts can only provide maps indicating territories, but not contain the territories themselves. The problem is how can this magic occur. This philosophic epistemological conundrum most
recently flashed into the science wars of the 1980s (for examples on both sides, Latour & Woolgar, 1979; Knorr-Cetina, 1981; Gilbert & Mulkay, 1984; Sokal & Bricmont, 1998; Gross & Levitt, 1994).

I propose here and in related studies (see for examples Bazerman, 1980, 1984, 1988), however, that we treat the distance between the world and the word not as a philosophical conundrum, but as a site for empirical investigation to see how scientists and other scholars manage the relationship and bring the world into representations (see also Goodwin, 1994). One useful research site to examine the complex tacit practical methods and methodological knowledge of disciplinary experts is neophyte learning, where practices and difficulties are more explicit and visible as students struggle to master them (see, for example, Bazerman & Self, 2017). Students are more likely to be aware of and talk about the processes they are still unsure of and expose by contrast what needs to be learned. In their learning and addressing new problems, they may also be in a position to propose innovations which they will have to defend.

The current study examines senior projects in political science at a mid-sized major research university in California. For context I interviewed seven faculty members for an hour each (audio recorded and transcribed) about the goals of the undergraduate program and their courses, writing assignments in their courses and the specific data requirements of those assignments. To study the research processes of students completing the senior project, I interviewed three students enrolled in the senior honors seminar that stretched across three quarters. The design was to interview every three times for an hour (audio recorded and transcribed). Once at the end of the first semester concerning their background in research methods and the use of data in previous assignments. Another at the end of the second semester concerning their processes and difficulties in data identification and collection, and finally near the end of the third semester after they had completed drafts of their projects and were making final revisions. I also collected drafts of their final papers. The three students who volunteered represented about a quarter of the cohort in the seminar and were offered gift card incentives for each interview they completed. They also reported enjoying and benefiting from the opportunity to talk about their work and become more aware of their methods and project challenges. Two of the students completed all parts of the study, but one withdrew before the third interview, so I have only the initial two interviews and no final paper for that one subject. The interviews for both faculty and three students were semi-structured.
The goals of this particular political science department for the undergraduate program were primarily to prepare critical citizens and politically engaged professionals, and not to prepare students for graduate studies or research careers. Even though the faculty themselves had substantial research careers and the department had a well-regarded graduate program, this view was articulated by the majority of faculty interviewees (though a couple felt it ought to be otherwise), often coupled with the observations that students who graduated from the majors would be going to a variety of careers. The few who would go to graduate school in the field would be trained in research methods at that point. In line with this philosophy, only about a dozen students each year, out of around nine hundred majors, were admitted to the senior year-long honors seminar where students were expected to develop and complete an independent research project.

There was only one required methods course in the major, and no additional methods courses were offered as electives. This course was a basic introduction to statistics, so students would be able to read and evaluate research literature, understand necessary manipulations and statistical logic, and then see how these methods would apply to political phenomena. While students were sometimes asked to carry out simple analyses based on provided data sets in order to understand statistical concepts, they were not asked to collect their data nor were they expected to be able to carry out independent statistical analyses at the end of the course.

While almost all courses had substantial writing requirements, the writing tasks tended to involve interpretation and evaluation of assigned texts, arguments weighing alternative positions of different political actors, strategy memos for organizations and campaigns using the theories of the course, and the like. Evidence and data where required were from secondary sources provided by the instructor. Thus, the few students who entered the senior honors seminar would have had little in the way of organized preparation in research methods, though as we shall see, they could have developed methods skills outside their coursework and then have applied the methods to the assignments, thereby transforming the nature of the tasks they carried out.

The three students whom I interviewed were quite distinct in their methodological preparation, concerning not only how to form empirical questions and design data gathering, but also how to analyze and bring the data into their arguments. In framing and carrying out their work for the senior project, the students depended on the methods repertoire they already had in place from
idiosyncratic individual experiences. To the extent they lacked experience with methods they could use for their theses, they faced challenges.

Student A. The first student had previously written only theoretical or conceptual papers in prior courses, requiring as evidence only quotations from theoretical sources. She chose quotations to use in her papers because she liked the idea expressed or could use it in her thinking. Although she had taken several history courses, these also did not require her to go beyond quotations from secondary sources. The theoretical ideas from her prior courses about electoral democracy led her to an empirical inquiry in her senior project about whether finance reform had increased democratic participation. Her initial approach to this question led her to try to correlate state legislation with donation and voting records before and after legislation. This data was available on government websites, but she needed to pull it out and organized it. She had never previously mined such data nor used Excel; accordingly, she had difficulties gathering the data and entering it on the spreadsheet.

The advisor had further suggested that the student’s topic required interviews to get a fuller picture from legislative proponents and opponents, those in charge of regulation, those who had taken issues to court, and media presenters. With guidance from her adviser, she focused on two states with similar strict laws, but different in administration. The student had never done interviews nor had ever designed interview protocols. She represented herself as having difficulties in identifying and enlisting interviewees and developing the interview guides. She had some difficulty in keeping herself to narrower, accomplishable goals and wanted to expand the inquiry to consider the effect of the Citizens United decision (expanding the rights of large donors to contribute to political campaigns) and to consider the perceptions of ordinary voters and other disempowered people. She also found her research question shifting and was unsure where it was headed. She would have to overcome many challenges to complete the study, as she admitted when asked about her difficulties:

The most difficult thing, and I think it’s pretty evident, is how hard it’s been for me to tie down to one research question, because I floated down for a while during the Fall Quarter and then Winter Quarter I started getting closer to what I wanted my research question to be, but it still wasn’t perfect. And the thing is, is that that is really driving all of my research, so when I get stuck doing my interview guides and I don’t know
why I’m so stuck, I think it ultimately comes back to these. I don’t know really what questions I want to be asking. I’ve never had to touch any type of public data to the extent that I’ve had now. Working with Excel sheets and formatting my own data, I’ve never had to do that either, I’ve never learned how to do that.

This is the student that did not respond to multiple requests for a final interview. I do not know whether she completed the thesis.

Student B. The second student also had a long-standing interest in political theory and had written a number of papers in which he elaborated, interpreted, and evaluated prior theorists, relying heavily on quotations. He also was interested in history and had a double major in history and political science. However, he had learned to gather and organize the ideas of theorists in Excel sheets, where he treated quotations as a kind of data to be examined systematically, rather than cherry-picked. He had first developed his use of spreadsheets for gathering data in two experiences as an undergraduate research assistant. For a history professor, he used Excel to collect and quantize patent rates in antebellum major US cities to compare regional industrial development. Then for a political science professor, he used Excel to categorize executive orders and actions by presidents over their years in office as part of quantitative analysis. As a result of his growing familiarity with Excel, he began using it for his papers, such as for his history thesis when he compared the integrity and consistency of the statements of Lincoln and Douglass in their famous series of debates in light of other statements made by them. There he organized their statements by topic in Excel. He expressed detailed interest and enthusiasm for the discoveries he made in each of these projects.

His political science honors thesis was more theoretical than the history one, proposing his own theory of the grounds and ends of liberal democratic government. In drawing on and analyzing the ideas of prior theorists, he again used Excel sheets to identify what theorists had to say on various issues. This allowed him to organize then and focus his discussion of each and to draw on them in support, challenge, or contrast to his own views. He also initially had proposed an empirical argument for the benefits of his theory, but that was dropped because it would present too much challenging work of a different sort within the time frame of the project. Accordingly, the thesis consisted of only the theoretical discussion of prior theorists and then the students’ proposals with conceptual justifications. The sources finally used
in the project were determined by what the student needed for the argument—that is helped establish pieces of the theory he was proposing. His proposal took up two-thirds of the sixty-page paper. This project helped the student elaborate and justify more deeply his previously held views, but by the student’s own admission, no views changed or new concepts adopted. When asked what he learned from this project, he talked about the reconfirmation of his ideas, though with improved clarity and justification. He did not mention any discoveries or changes in thinking:

I have sharpened significantly the ability to slow down, to progress with my argument and challenge myself every step of the way. [My Adviser] was critiquing me so much on each of these individuals I was almost doubting myself…. Pulling out and teasing out the arguments and the implications I realized, yeah, I was absolutely right, but now I have all the evidence in the world to show it. That was immensely beneficial.

This thesis then is an elaboration of previously held views based on postulates, but growing from the author’s perception of social problems and contradictions in traditions of Western thought. The primary systematic evidence is in theorist quotes initially collected in a spreadsheet but then discussed in a more discursive form. This student used one data collection method throughout all projects, recording in spreadsheets information from textual resources, whether patent lists or theoretical texts, organizing the entries, and then analyzing or discussing the material in qualitative or quantitative terms. He did not venture beyond that for collection, recording, or analysis. While this method had led him to discoveries in earlier projects, here he used it only for elaboration and reconfirmation of prior beliefs.

Student C. The third student had the strongest background in empirical methods and kept extending them as her studies permitted. In high school, she had taken an Advanced Placement course in Statistics and when she took the methods course in the major, she could focus more on how political science used quantitative data to answer disciplinary questions rather than on the basic statistical methods. She had developed interviewing skills in some projects and had used both quantitative and qualitative data in many of her papers in university courses. For example, in an environmental policy class, she collected a variety of kinds of evidence to determine the positions of the actors in a city debate over the construction of a desalinization plant.
Her data included city council minutes, news reports and presentations of activists, and bureaucratic regulations.

Further, she examined each of these for their reliability and confirmation of her judgments. She repeatedly expressed in her interviews concern for selection and evaluation of data for analysis in all her projects. She also reported reading doctoral level methodology books on her own.

For her thesis she took on a methodologically complex project, incorporating many kinds of quantitative and qualitative data, public records, and interviews. To evaluate the effects of educational policy on professional development implemented in school, she had to understand and gather evidence on the detailed institutional mechanisms of implementation as well as the complexities and obscurities of school system budgets. Further, she made sophisticated decisions in studying most-similar comparisons and paying attention to the often-ignored median schools as well as high- and low-performing ones. Even as she was in the middle of the project, she was aware of the importance of her methodological reasoning and choices. In the second interview, while she was in the middle of the research process, she commented:

My methods chapter has been the hardest to write but it’s been very beneficial to me in my head, organizing what I need my evidence to do. ...So, it is different for me to have to defend every methodological decision I have made—I’m doing this because of this and this because of this.

This student was able to evaluate the quality of the data from each source and to use different kinds of data to triangulate and to dig more deeply into each. She was specific in what she intended to get from interviews teachers, principals, and staff, and she produced well-structured interview guides. She then used clues gleaned in the interviews to locate and interpret policy and budget documents. Even as she dug into the details of the project, she was aware of the need to keep her eye on larger issues and did so even as she kept narrowing the focus of the data she was using. Theory shaped her research and indicated the kinds of data that might exist, but did not prejudge what she would find or what documents might actually exist or say. While she was conscious of the limitations of her skills and the study, the student made many discoveries revealing how complex the processes were that she needed to examine in order to evaluate the effect of particular policy decisions in
their various implementations in different jurisdictions. When asked what she learned, she commented:

Oh my gosh. I have a lot. I think the biggest thing I've learned is that research takes forever... I think my question was harder than a lot of their [classmates'] topics, [but] I think my topic is a lot more answerable.

Her care, attention, and patience in thinking through, refining, and carry out her methods, based on her previously developed methodological sophistication, resulted in her honing her questions and finding empirical ways of answering them with precisely relevant data. Her understanding of her methods allowed her to know what kinds of questions she could answer, how to specify the questions as she gathered her data, and to pin down the data that would provide the answers. This thoughtful interaction between question and methods made her questions answerable and led to discoveries. She came up with novel findings that revealed how policies were implemented and how administrative arrangements and cultures at the school and district level influenced the experience of teachers.

When we compare the three student cases we can found that Student A’s lack of experience in empirical methods created difficult (perhaps insurmountable) problems in focusing her question and developing data for her thesis, and it is unclear what discoveries or substantive learning she was able to gain in the course of her project. Student B did have a dominant method of gathering, recording, and organizing data from existing documentary sources developed over multiple experiences. This method enabled a particular style of inquiry, but restricted bringing in or recording more extensive sources of data and led to the triaging of a proposed empirical argument for the benefits of his theory. The data recording method became in the thesis a device for confirming and elaborating prior beliefs, placing them within the matrix of the beliefs of prior theorists that could be readily collected in the spreadsheet. Student C, with broader and more flexible methods resources and greater methodological reflexivity, was able to draw on a wide range of methods, to make numerous thoughtful choices concerning data sites to examine, to identify ways of locating and collecting data, to use data sources to locate other data sources, and to analyze the data complexly. The methods were used as tools of discovery rather than idea confirmation devices. Accordingly, the thesis provided empirical answers to a focused question and
contained discoveries about the way the policies were implemented and the reasons for the differences of implementation.

Overall the student interviews suggest that prior training and experience in gathering and manipulating data affected numerous parts of the thesis writing process, with consequences for the final thesis. Some of the aspects influenced by differences in methods and methodological experiences include the nature of the inquiry and formation of the research question, the understanding of data needed to deliver on the question, flexibility and variety of data gathering methods conceived and deployed, precision of implementation, sophistication of learning while doing the project, character of discovery made in the project, interpretation of data, and understanding of complexity of phenomena investigated. Thus, methodological sophistication may contribute to writing’s role in cognitive development as students struggle with the challenges of developing coherent accounts of the world represented in their data. This can push students to go beyond repetition of what they have learned from their readings (see Bazerman, 2009 for a discussion of the relation of writing to learn and writing for cognitive development). All these differences added up to differences in the overall character and quality of the argument made in the thesis. These initial observations suggest a link between the quality of writing and the training and experience in research methods—and thus a link between methodological development and writing development, at least in those domains that rely on empirical development.

In this particular instance, the prior learning of methods and development of methodological sophistication were not primarily the result of an organized curriculum but were based on idiosyncratic individual experiences. The idiosyncrasy of experience in this study heightened the differences in this study, thereby making more visible the relationship between methodological experience and writing practices, processes, and results. The results suggest that academic departments, in order to prepare students for successful empirical projects, may want to be more intentional in the sequencing of writing tasks requiring data gathering and use across their course sequences. Writing education also should attune students to the importance of the quality of data and data use in their writing, as well as to help students understand the reasons for and underlying logic of different disciplinary methods. Further students would benefit from being introduced to how these differences of methods are related to the forms of disciplinary writing and argument, even as some of the particulars must be left to experiences embedded within the disciplinary courses.
Most students have the intuitive sense that good writing requires good stuff to write about, but what they may not be aware of is how important is the process of identifying, collecting, recording and analyzing the stuff is for their writing. Even more, they may lack the tools and skills to be able to engage in these data-related processes that will give them the good stuff to write about. Both writing programs and academic programs in the disciplines may want to consider their responsibility for providing the experiences that will enable students to identify and represent the good stuff to make high-quality academic arguments.

References


