

1-1-2002

A Methodology for Streamlining Historical Research: The Analysis of Technical and Scientific Publications

John T. Battalio
Boise State University

A Methodology for Streamlining Historical Research: The Analysis of Technical and Scientific Publications

Abstract—This article provides a framework for organizing and structuring the research of historical researchers who analyze technical and scientific publications. Because historical research spans both decades and centuries, an effective research methodology is essential. The framework consists of a multifaceted 10-step method for studying the written discourse of scientific and technical communication, specifically for interpreting historical data obtained from articles published in technical and scientific journals. The method is a reliable means for making sense of the enormous body of data that awaits historical researchers in the volumes of scientific and technical discourse already published.

—JOHN T. BATTALIO

Index Terms—Analyzing technical and scientific publications, historical research, research methodology, scientific discourse study.

Manuscript received May 30, 2000; revised September 24, 2001. The author is with the Technical Communication Program, Department of English, Boise State University, Boise, ID 83725-1525 USA (email: j battali@boisestate.edu).
IEEE PII S 0361-1434(02)02327-5.

In a recent article in IEEE TRANSACTIONS ON PROFESSIONAL COMMUNICATION, Grove and Zimmerman [1] describe the emergence of technical communication as an academic discipline. However, the field cannot yet be considered a full-fledged profession, as noted by recent studies examining the field's professionalization [2] and the availability of research-methods course work for technical-communication students [3]. Grove and Zimmerman caution that, in order for the field to have credibility, it needs frameworks for organizing and structuring its "bodies of knowledge and ... [guiding] subsequent research" [1, p. 158]. Such frameworks enable macro views of the field and provide "a solid foundation for future researchers to advance the field" [1, p. 158]. To create a framework for establishing

technical communication as a legitimate academic discipline, I propose the multifaceted 10-step method described here as one method for systematically collecting historical information from a study of scientific and technical journals. Clearly, the history of technical and scientific communication may also be researched by analyzing other forms of discourse and other time periods before the advent of journals. Tebeaux and Killingsworth [4] and Connor [5] describe methodologies for developing historiographies of individual and collective scientific and technical texts, including those before the advent of professional journals. However, the focus here is specifically on the analysis of journal publications.

Scholars from a variety of areas have studied the historical impact of scientific and technical writing,

including professional journals. One such group of scholars has already begun the arduous task of tracing the historical origins of science and technology. Moran [6] provides an overview of this research; Kynell and Moran [7] contribute a more recent update. Tebeaux [8]–[10], for instance, has researched technical communication in medieval and Renaissance England. Others have described historical contributions by women [11], [12].

Genre studies have also played a prominent role in describing and contextualizing the scientific and technical journal as genre. Bazerman [13] traces the development of the journal article in the field of physics. The research by Berkenkotter and Huckin [14] of 12 scientific journals expands Bazerman's study. Swales [15] analyzes the various parts of the research article. Atkinson [16] contributes sociohistorical analytical methodology in his study of the *Philosophical Transactions of the Royal Society of London*. Finally, Orr [17], as part of IEEE TRANSACTIONS ON PROFESSIONAL COMMUNICATION Special Issue on Engineering Genre, has described genre in the fields of computer science and computer engineering.

Finally, numerous studies of the emergence of disciplines have traced the evolution of a variety of scientific and technical fields by studying their writing: geology [18], experimental parapsychology [19], sociology [20], astronomy [21], physics [13], biology [22], ornithology [23], and occupational therapy [24]. In general, these studies attempt to show the relationship between the evolution of the field's writing and the professionalization of the field itself. The common methodology compares changes in content and organization, in graphical presentation, in parts of the research article and/or in linguistic, grammatical, and syntactic structures with the

evolution of the field professionally. Such professionalization may be evidenced by factors such as the appearance of professional associations and journal publication, methodological sophistication, membership dynamics, contributions to theoretical knowledge, prominence and influence of individual members, attainment of professional acceptance, and exclusionary forces, that is, of expert from amateur. Analysis may involve linguistic, rhetorical, sociological, and/or philosophical approaches that include either quantitative or qualitative methods or both.

One avenue for such historical research is an analysis of the data obtained from articles published in technical and scientific journals. Because a historical search spans both decades and centuries and can include many volumes of articles, an efficient research methodology is essential in facilitating the process. My purpose here is to propose a flexible methodology that researchers can incorporate into their historical studies of the written discourse of scientific and technical communication. Specifically, I will do the following:

- suggest a rich, multimethod research design that uses a specific sampling methodology and selection of variables;
- recommend a time continuum useful for the research of scientific discourse;
- suggest a number of ways, both formally and informally, quantitatively and qualitatively, to use the numerical data obtained from this methodology.

Methodologies for empirical research in the technical communication and related literature may begin with qualitative analyses as starting points for identifying research questions (see, for instance, descriptions by Goubil-Gambrel

[25, pp. 587–589], Grove and Zimmerman [1, p. 162], Lauer and Asher [26, p. 45], Sullivan and Spilka [27, p. 596], and Thompson [28, p. 51]). Other methodologies have emphasized quantitative approaches (see, for instance, descriptions by Goubil-Gambrel [25, pp. 585–587], Lauer and Asher [26, p. 82], MacNealy [29], and Spyridakis [30]).

Other researchers favor multiple approaches. In fact, Charney [31] points out that quantitative and qualitative methodologies are most effective when viewed as complementary or even overlapping. Qualitative analysis is described by Sullivan and Spilka [27] as a complex and versatile methodology that can be implemented throughout the research process. Berkenkotter and Huckin [14], for instance, combine interview methodology with a quantitative analysis of a systematic sampling of scientific journal articles. Quantitative studies are also compatible with alternative research methodologies. For instance, Atkinson [16] integrates rhetorical analysis with multidimensional register analysis, and Longo [32] juxtaposes quantitative research and cultural studies. Discussing his research methodology, Bazerman echoes this interrelationship when he notes that “the statistics are to indicate **that** something is happening and the close readings are to find out **what** that something is” [13, p. 159, my emphasis]. In particular, the value of both methodologies has already been described by Charney [31] and Markel [33].

Similarly, the methodology described here may be viewed as a rich, multimethod design that historical researchers favoring numerous methodologies may incorporate into their research. This approach outlines a methodology where 1) numerical data is gathered from a large body of discourse; 2) data is then plotted

in various ways in order to narrow that large body of discourse into a much smaller subsection for further research; and 3) if the data warrant further analysis, that subsection of discourse is then studied in-depth. In this way, historical researchers may sample a very large body of discourse in a relatively short time and then, after quantitatively analyzing the results, continue their research in a number of ways: 1) focus qualitatively on specific time periods, authors, movements, and other areas of interest uncovered by the data; 2) continue with further quantitative analysis; or 3) implement other methodologies such as critical research [34]. Thus, the methodology described below is a multifaceted means for analyzing a large body of historical data.

As a result, this methodology may provide a significant savings of time for both qualitative and quantitative researchers. Take, for example, a researcher studying the publication history of a field, the journals of which span the last 150 years. A qualitative approach may involve interviewing noted contemporary authors and surveying the profession's active members. However, this approach will not be possible for the first hundred years because subjects living in this time period are deceased. Thus, a determination must be made of the significant figures, events, and movements of the time. Biographies, autobiographies, and professional and personal correspondence must be discovered and analyzed to approximate the results not obtainable through interviewing and surveys. For quantitative researchers studying the same database, coding and interpreting enormous sets of data from potentially thousands of articles is a daunting task. On the other hand, if a means can be found to pinpoint a critical time span of the historical topic under study, historical researchers may then be much more productive. They could

customize their research, whether qualitative or quantitative, to a more limited span of time using the set of procedures described below.

The method I suggest is not intended to invalidate other methodologies. It may be used on its own or may complement and facilitate current methods. Because the methodology may be integrated into either qualitative or quantitative studies, it provides the complementary relationship between the two methodologies that Charney [31] and Markel [33] describe.

THE METHODOLOGY

The proposed method outlined below moves from a quantitative to a qualitative approach. Those who focus on qualitative research may wish to implement the methodology when a large body of potentially useful data is encountered. At that point, the following methodology will aid in narrowing down the database for further qualitative analysis. Finally, steps one through nine may comprise the procedure for a quantitative approach. Fig. 1 shows the following sequence of steps.

1. Determine the focus of the historical study by developing a specific research question or hypothesis, selecting the relevant journal(s) in the field, and determining the appropriate number and kinds of variables and factors to be used in analyzing the count data generated by the study.
2. Search for overviews of the profession, provided, if possible, by the field's own scientific or technical community.
3. Use systematic sampling and select a database of articles for analysis.
4. Read the articles, placing each one in a historical

data matrix representing the historical continuum discovered, or created, in step 2.

5. Plot the data from step 4 in order to look for historical patterns and to refine the factors and levels of the data.
6. Identify changes in the discourse over time, such as in the use of graphics, formatting, and citations in the cross section of articles.
7. Analyze a smaller sample of the original body of articles.
8. Use categorical data analysis on the numerical data obtained from the study.
9. Develop discourse schemas.
10. Implement qualitative analysis.

It should be noted here that a related methodology is content analysis. See Berelson [35] for the standard description and de Sola Pool [36] for implementation. Krippendorf [37] and Weber [38] offer more recent descriptions. I describe the methodology in step eight below.

Briefly, content analysis gathers inferences from texts by means of quantitative and qualitative methods. The methodology is useful in historical studies since it allows the researcher to infer events from texts that are located in their appropriate historical contexts [37, p. 24]; see also Weber [38] for numerous examples.

However, unlike content analysis, the purpose of the methodology described here is to narrow down a large body of information in order to focus on a smaller, more informative and more manageable body of data. Qualitative content analysis, which I mention in step ten, may be used for this latter set of data. Likewise, individual components of the methodology may employ quantitative content analysis. For instance, step seven's linguistic methodology is akin

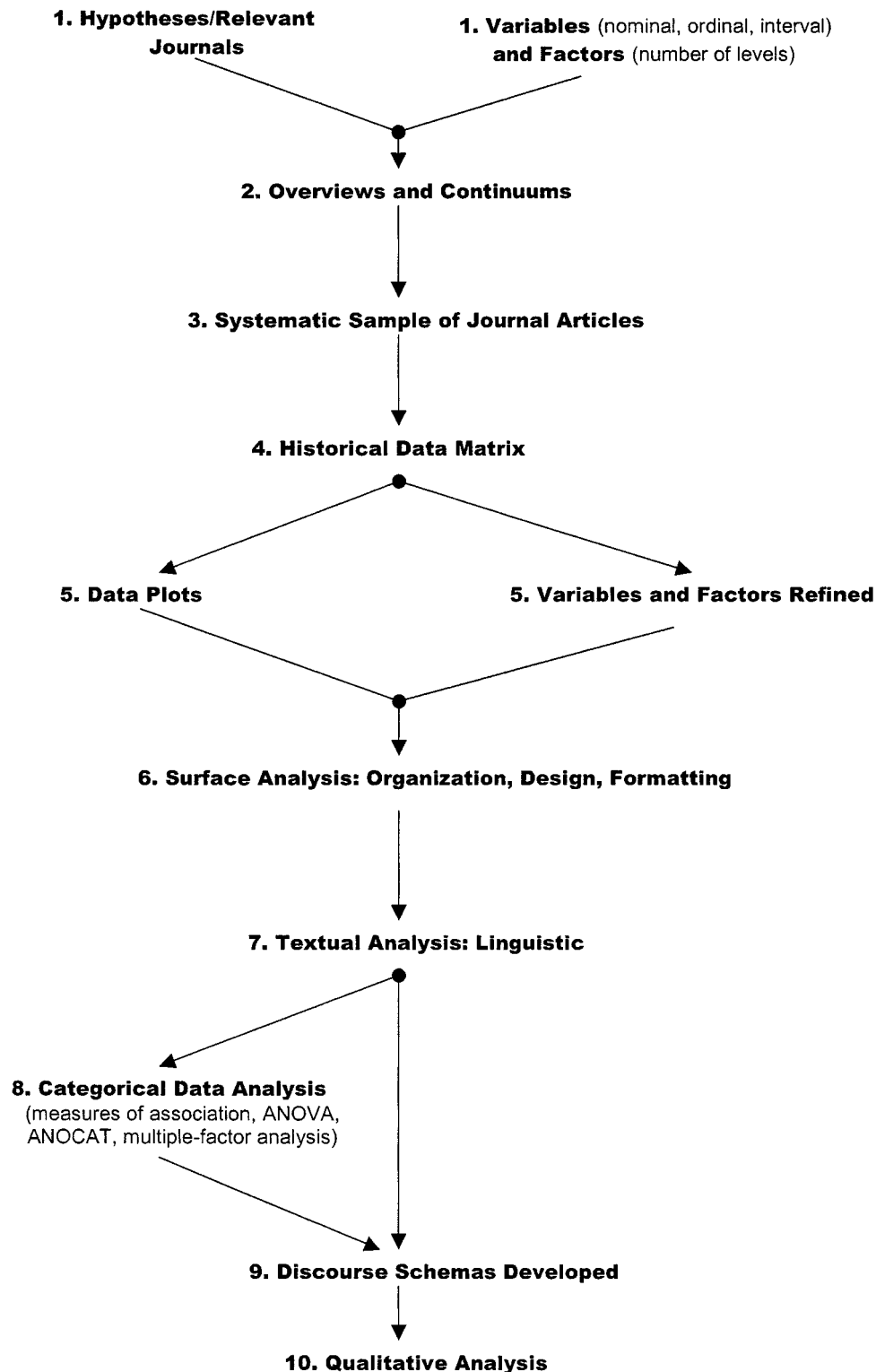
to content analyses based on interpretation of indices [37, p. 40]. Content analysis may also be used to prepare data for the categorical data analysis described in step eight. Weber [38] best

describes computer applications for this purpose.

I have used the methodology described here in a recent study of the professional literature in the

field of ornithology (Battalio [23]). Since some of the results of my research parallel already known features of scientific discourse, I believe that this method is one reliable means of making sense

Fig. 1. Overview of the methodology.



of the enormous body of data that historical researchers may uncover in the volumes of scientific and technical discourse already published.

The following sections describe each of the steps in the methodology.

1. Determine the Focus of the Study As stated above, this methodology is targeted for studies that trace the historical origins and development of a technical or scientific field by analyzing its publications. The analysis should begin with a problem that needs solving or a research question that needs answering. In some methodologies, research questions may be refined or reformulated during the ongoing research process as more data are uncovered. The research may then follow different paths. In other methodologies, null hypotheses are first stated in terms of what the data would normally be expected to show. Alternative research hypotheses are then formulated in terms of what would not be expected. If the data from the research study show unexpected results, then the null hypothesis is rejected in favor of the alternative one; and the research results may be viewed as noteworthy. Miller [39], Lauer and Asher [26], MacNealy [29], and Ott [40] describe hypothesis building in more detail.

The researcher must then answer two critical questions before beginning the actual reading of texts: 1) What journal or journals will be the focus of the study? 2) How will the data be classified for later analysis?

Journal Selection: An informative snapshot of the evolution of a profession's writing may be obtained if a key journal in the field is selected. Whether or not the main journal in the field is chosen depends upon the focus of the study. The major journal should be chosen if the

researcher wants to study the major argumentation and forces that influence the field. On the other hand, if there is an interest in a broader base of information, then a publication that appeals to a general audience may be more informative. For instance, some fields publish semiprofessional journals, general magazines, or newsletters for a diverse readership. Studies that intend to compare amateur-professional discourse, for instance, will include data from these latter publications. The mere presence of such publications will indicate the amount of specialization the field has attained, the extent to which the field appeals to a wider audience, and the composition of those audiences.

In order to understand the full development of the professional field, the researcher should select a journal with a lifespan paralleling that of the field to be studied. Two options are available. One option would be to begin the study from the journal's origins. This method would be useful for studies of the professionalization of technical and scientific fields or of the interaction between amateur and professional. Numerous scholars have identified characteristics of professionalization, including standardized formats, established research agendas, use and volume of citations, linguistic variation, and the presence and complexity of data analysis and graphics. Although most scholars agree that science, for instance, had professionalized by the end of the 19th century, some fields professionalized much later. For instance, data collected and analyzed over the life of an ornithological journal indicate that American ornithology did not professionalize until the mid-20th century [23].

Another option, primarily for those wishing to research developing paradigms or theoretical positions within techno-science, is to research a specific point in time

in the journal's history. For fields with very diverse and wide-ranging research agendas, researchers may choose articles that focus on a specific research topic over an extended time period. This specificity would help standardize the data by limiting variations due to changing focus, methodology, and researchers, as well as by providing a more standardized technical vocabulary for analysis. Bazerman [13], for instance, successfully uses this method in his analysis of physicists' discourse by restricting his study to articles dealing with the study of spectroscopy. Anthony's study evaluating a standard model for describing the structure of research-article introductions restricts article selection to those articles in IEEE TRANSACTIONS ON SOFTWARE ENGINEERING receiving awards in 1996 [41].

Journal selection can be more complicated, however, in the case of journals that abruptly change their formats and/or objectives. For instance, in 1992 the Association of Teachers of Technical Writing abandoned *Technical Writing Teacher Quarterly*. Editors Wahlstrom and Lay attributed the changing nature of the field of technical writing as the reason for the move from a strictly pedagogical journal to one that "integrate[s] pedagogy with solid research and theoretical studies" [42, p. 6]. In a similar move, in 1996, the Society for Technical Communication (STC) refocused its journal *Technical Communication*. According to newly appointed editor Hayhoe, recommendations by a panel established by the STC board of directors prompted the changes, the intent of which were to make the journal "universally acknowledged as the preeminent journal in our field" [43, p. 126]. Peer reviewing all articles, implementing an editorial review process, and moving practitioners' articles to the monthly *Intercom* magazine were key changes. Since

changes such as those occurring with these two journals likely represent paradigm shifts within the discipline, these events in themselves prompt further study. However, for historical studies over the life of a journal, major shifts in focus may cause such radical changes in format, content, and contributors that historical studies of each variation of the journal may best be conducted separately.

Data Classification: The second question to be answered concerns the coding of data. There are three main types of data, or variables of interest: in order of flexibility, NOMINAL DATA, in which ordering is not important (for instance, the sex or race of journal authors); ORDINAL DATA, in which order is clear, but absolute distances between each are not (for instance, a journal's level of grammatical complexity such as use of simple, compound, and complex clauses); and INTERVAL DATA, in which data is ordered and quantifiable (for instance, membership figures and the number of graphics used during a journal's history).

Because count, or categorical, data analysis (described in step eight) uses data tables, or matrices, as a basis for its calculations, the study is more effective if the variables, or factors, of interest described above can be selected based upon their ability to be coded into multiple levels, as for instance, levels of graphics as for lay, amateur, and professional reader; levels of sentence structure as simple, moderate, and complex; levels of changes in argumentation as methodological, theoretical, and model-oriented; and so forth.

Throughout the collection process, the researcher should look for additional data categories that sometimes become evident only during the actual accumulation of data. If the study can be designed so that at least one variable, or factor, has only two levels and is a distinct response variable (that is, a dependent variable upon which

the other variables operate), then more sophisticated statistics as described in step eight may be implemented. For instance, in a hypothetical study, a researcher is studying the professionalization of a technical field by analyzing a database of articles from the field's journals. S/he has determined an approximate date of professionalization and has hypothesized that a number of factors, or independent variables, such as sentence length, changes in graphical presentation, article length, and professional collaboration can be used to identify the time span within which the field has professionalized. Numerous researchers use such empirical evidence to evaluate the maturation of professional fields [13], [18]–[21], [23], [24].

Thus, the dependent variable, professionalization, has two levels—yes or no. The independent variables have varying levels depending on the kinds of data collected. The researcher would then compile a table for each independent variable. For example, if the researcher had grouped professional collaboration into three levels of authorship of 1, 2–5, and 6+ authors, the table would appear as shown in Table I. That is, in row one, of the 71 single-authored articles, 64 were written prior to the determined date of professionalization; 7 were written afterward, and so forth.

If this level of data categorization is not possible, categorical data analysis still provides useful descriptive information.

The reliance on levels of factors do point to a disadvantage of categorical data analysis, since it is ordinarily better statistically to collect continuous data through simple random sampling in order that ANOVA and regression methodologies can be implemented. Of its nature, categorical data analysis, which relies on grouping individual data values within the levels of the factors being analyzed, does cause some loss of information which creates a need to supplement statistical findings with qualitative historical research. If nothing else, statistics can at least direct the researcher toward the most useful factors and data for further study, as described in step ten below.

2. Search for a Historical Overview

A historical overview of the field or research topic is an invaluable aid in data collection and analysis. Such an overview not only directs the researcher to specific time periods of interest, but it also parallels the establishment of statistical norms. Acquiring a historical overview may be likened to determining benchmark values, as in formulas for determining the number of observations needed for random sampling whereby the accuracy of the results (in t or Z table values), the variability of the sample (σ or σ^2), and the width of the interval (E) are calculated. Two common methods for determining these statistical norms are to review the statistical literature for commonly accepted values or, if none exist, to do a pilot study to obtain these values. Similar methods may be used by the historical researcher: reviews of the literature and interviews with key figures provide historical context; pilot studies may also be conducted.

Reviews of the Literature: Researchers of technical and scientific discourse should first perform a literature review of the target field itself, of related historical fields and of journals focusing on discourse studies

TABLE I
SAMPLE DATA TABLE SHOWING
PROFESSIONALIZATION STATUS OF
AUTHORS IN SAMPLE OF ARTICLES

| No. of Authors | No | Yes | Total |
|----------------|-----|-----|-------|
| 1 | 64 | 7 | 71 |
| 2-5 | 31 | 97 | 128 |
| 6+ | 5 | 27 | 32 |
| Total | 100 | 131 | 231 |

in order to obtain a historical overview. Connor [5] offers a useful framework to provide historical context for studies of the history of technical communication. Of special note are several key steps in her methodology. Researchers should review specialist historical literature, keeping in mind the interpretive nature of any historical work. They should also determine the journals' intellectual and social context by referring to works by professional historians and scholarly publications on the history of the time period under study. Understanding the journal's readership is another critical contextual component, which the journal itself should provide in the form of features such as Letters to the Editor, Commentaries, Short Communications; yearly membership figures; minutes of annual meetings; summaries of annual conference proceedings; and notices of awards, special events, and calls for papers.

A survey of the main journals in the field should uncover at least a few historical articles. For instance, Abt [21] has written about astronomy, the Physics Survey Committee [44] has researched the field of physics, and ornithologists have published occasional historical overviews [45], [46].

A helpful comparative benchmark would be a chronological overview provided by the profession itself. In particular, a continuum of activity is a valuable aid to the historical researcher. Researchers studying field and laboratory sciences will find extremely useful the continuum of perspectives offered by Killingsworth and Palmer [47, p. 112]. The continuum shows that the progression from natural history study to experimental scientific research may be viewed as a continuum from simply observing, to observing and measuring in the field, to manipulating in the field, to manipulating in seminatural conditions, to bringing the system

into the lab, and finally, to manipulating theoretically using mathematics and computer models.

Interviewing: Interviews with key figures in the target field, such as prominent researchers or practitioners, and with noted historians of the field or periods under study also assist a historical overview. For instance, Killingsworth and Palmer [47] developed their continuum based on interviews conducted with ecologists at then Memphis State University to describe the progression from natural history study to experimental scientific research.

Pilot Studies: If the researcher is unable to find appropriate overviews from outside sources, then a pilot study should be conducted in order to obtain a historical overview of the field of interest. The study might consist of a survey of a small systematic sample of articles, say the first article of a representative journal from each issue for a given number of years. The exact nature of the study would determine whether the specific analysis of the sample would focus on methodological, linguistic, or theoretical changes over time. For instance, for a study of technological change, a count might be made of the various technologies described in the sample of articles over the time period under study. The data from this count would give the researcher an overview of the number and kinds of technologies influencing the technical field and the evolution of these technologies over time. Pilot studies such as these are useful at any point in the research process where baseline values or statistical norms are needed.

3. Select a Database for Study A major difficulty prompting the use of the research methodology described here is the large body of information the historical researcher may need to digest.

For instance, a complete analysis of three hundred years of journal articles is prohibitive. Another difficulty is the inherent bias associated with discourse accumulated over large time spans, such as a variety of research methodologies, individual writing styles, influence of journal editors, repetitive authorship, and changes in research agendas. A random sampling of articles would enable the researcher to obtain a more manageable sample and manage such bias, as well as allow for statistical analysis. If circumstances such as those listed above do not exist, however, then random sampling may not be necessary.

One type of random sample, a systematic sample, is particularly useful for historical research in that it provides a cross section of data. It adds a historical perspective in addition to providing a random sample for investigating statistical inferences. Its value for historical researchers is that it allows the researcher to incorporate a chronological component into the database since the sample is spread more evenly over the population [48, p. 205].

Systematic sampling involves selecting one unit at random from each stratum, then every k th unit thereafter. For instance, in a study of articles over a hundred-year span in which the journal has published five issues yearly, the unit is the issue and the year is the stratum. Therefore, there are five units per stratum. The researcher may roll a dice to determine the issue number to be sampled; that is, if a one were rolled, then the first issue of each year will be used. The researcher may then use another method to determine which article(s) in that issue will be selected. For example, if the purpose of the study is to determine the changes in research topics over the hundred-year period, the first article of the issue may be chosen, since placement of the article

will indicate the research article's value. This procedure will provide a 100-article database by using the first article of the first issue each year for the hundred years under study. If a larger database is desired, more units and/or articles per unit may be chosen.

As long as this grouping of articles into one-year increments does not introduce bias into the sample as described above, systematic sampling is easy to implement and will likely give more precise results than simple random sampling [48, p. 205] because it more uniformly provides a cross section of samples and is more selective. In addition, systematic sampling may approximate stratified and cluster sampling, thereby widening its applicability as well.

Another issue to be considered when selecting a database whose data may be used statistically is sample size. For normally distributed data, statisticians have devised a number of statistical formulas for estimating an optimal sample size. However, since count data is not normally distributed, these formulas are not as useful. A more important issue in analyzing categorical data statistically, as described in step eight, is the individual cell-frequency count, which may assist the researcher in evaluating the adequacy of the sample size. As a rule of thumb, counts for individual cells in the data table should total at least five. Alternatively, Agresti [49, p. 246] suggests that the total sample size divided by the number of table cells should be greater than five. The data in Table I, for instance, may be analyzed by categorical methods because all cells have at least five occurrences and the total sample size (231) divided by the number of cells (6) is greater than five (38.5). If this were not the case, then the researcher would continue sampling using the established methodology until these minimal values were obtained. If the resulting tables were to include cells with

unusually large values, then other methods, like combining cells or collapsing tables, would be preferable. Agresti [49, p. 239–250] suggests a number of alternatives. Battle et al. [50] provide formulas for using statistical power to determine sample sizes.

4. Read Chronologically the Cross Section of Articles The researcher now has the tools necessary to begin a historical analysis: a carefully selected journal from the field, a historical continuum, and a random sample of articles to be read and analyzed in light of the historical continuum. Next, the researcher reads each article in the database, categorizes each according to the elements under study, and then places each element or article along the historical continuum. The resulting data table may then be used for analysis.

Mention should be made of the need for internal consistency whenever data are categorized or coded. If the categorizing or coding of data is interpretive or dependent upon varying degrees of subjectivity, two or more raters should be used for this part of the data collection process. Raters must first practice categorizing or coding sample data until they achieve a high level of agreement, or correlation. During the actual data collection process, the level of agreement among the raters is recorded and reported, ordinarily in the form of reliability coefficients. These correlation values range from -1.00 to 0 to $+1.00$, with $+1.00$ being perfect agreement. Minimum levels of agreement and reliability coefficient values vary among methodologies. Krippendorff [37, pp. 72–74, 146–154] and Lauer and Asher [26, pp. 138–139] discuss inter-rater reliability in more detail.

A specific example may best describe the methodology presented thus far. In order to clarify the process of

transition from natural history to experimental science as American ornithology professionalized over the past century [23], I used *The Auk*, American ornithology's professional journal, as a historical barometer of this progression. The database consisted of all major research articles (444 in all) from all four issues of 11 volumes of *The Auk* beginning in 1890 and each 10-year interval thereafter; i.e., 1900, 1910, etc. After determining the research topic of each of these articles, the resulting data were compared with the historical overview provided by Killingsworth and Palmer's Continuum, as described in step 2 above. The merger of data and overview clarified the progression of ornithologists' writing from natural history to experimental science discourse and thus showed the evolution of their writing during the past century. Table II shows the distribution of articles across the continuum. Because the systematic sample distributed the articles over *The Auk's* 109-year history, the table shows a representative cross section of topics of interest to ornithologists over the years and reflects methodology in keeping with what was considered ornithological science at each respective decade.

5. Identify Historical Patterns by Plotting the Data Once the data are gathered, the researcher should obtain an overview of the entire set of raw data because a preliminary look at this data may help to determine which components of the data will be useful for further analysis. Descriptive statistics, whereby the complete set of data is simply plotted in tables and graphs, provides such an overview. In addition to the count-data matrices (see step 1) and data tables (see step 4), graphs also provide insight into newly obtained numerical data. Line graphs are especially insightful for historical research since a major focus is to identify changes over time.

Such graphs can be easily created using the data table from step 4 since a systematic sample has been used as the database and the use of the historical continuum has organized the data chronologically. The researcher may use a computer spreadsheet program, like Microsoft Excel, to draw line graphs (or bar graphs) that plot time along the horizontal axis in order to search for patterns of historical change. Such plots are also useful in determining the kinds of inferential statistical procedures to be used in analyzing the raw data (see step eight), as well as in refining the kinds of variables and levels of factors used in the study.

Three examples show the value of plotting historical time lines. In a study of the probability distribution of the age of references in engineering papers, Matricciani [51] uses the graphic shown in Fig. 2 to depict the age of references cited in an Italian electronics journal during a 60-year period. The line graph plots the number of references in quartiles between 1930 and 1990. In my study of ornithologists'

discourse, Table II shows that the bulk of the data from the 444 articles occurs on either extreme of the continuum. Bar-graphing this data and adding a mean line to reveal the change over time, as well as the variation in the data, clearly demonstrate not only the distinct change in topics over time, but also the four periods of ornithological discourse. The patterns derived from these plots clarified the scientific community's response to the articles under study. Finally, Paul's [52] study of the rhetorical use of citations includes several plots of the occurrences of citations over a 17-year period. As shown by these three examples, plotting data facilitates the researcher's ability to see patterns in data otherwise obscured by raw numbers.

6. Identify Changes in the Discourse In order to complement the historical overview, researchers may also gather other count data. As described in step 4 above, multiple raters should also be used here if appropriate. Frequency and types of graphics and the amount of coauthorship, as well as changes in organization, formatting, and article length

may also be quantified. The appearance of standard features, such as the IMRaD format (Introduction, Methods, Results, and Discussion) and the presence of acknowledgment sections and literature reviews may likewise be of significance. For example, Abt [21] found to be useful a plot of the lengths of articles published between 1910 and 1980 in the three major American astronomical journals. The graph shows the striking increase in paper lengths since 1940, as shown in Fig. 3.

It should be noted that any direct comparison of data over multiple graphs should be made only after the data have been standardized. For all graphs being compared, their horizontal and vertical axes should show the same kinds of data and use the same incremental values.

Alternatively, the researcher may collect all of the data at once by combining this and the following step with steps four and five that first provide a historical overview. However, a researcher who first acquires a chronological overview of the field's historical development

TABLE II
ORNITHOLOGICAL DISCOURSE IN *THE AUK*: 1890-1990
[NUMBER OF ARTICLES (PERCENT)]

| Year | Observe in Field | Observe & Measure in Field | Manipulate in Field | Manipulate in Semi-natural Conditions | Bring System into Lab | Theor. Sys./Math & Comp. Models | Other | Total No. of Articles |
|--------------|------------------|----------------------------|---------------------|---------------------------------------|-----------------------|---------------------------------|-----------|-----------------------|
| 1890 | 17 (41%) | 6 (14%) | 0 (0%) | 0 (0%) | 17 (41%) | 1 (2%) | 1 (2%) | 42 |
| 1900 | 14 (40%) | 2 (6%) | 0 (0%) | 0 (0%) | 16 (46%) | 1 (3%) | 2 (6%) | 35 |
| 1910 | 24 (69%) | 6 (17%) | 0 (0%) | 0 (0%) | 2 (6%) | 0 (0%) | 3 (9%) | 35 |
| 1920 | 20 (61%) | 2 (6%) | 0 (0%) | 0 (0%) | 8 (24%) | 1 (3%) | 2 (6%) | 33 |
| 1930 | 20 (50%) | 5 (13%) | 0 (0%) | 0 (0%) | 7 (18%) | 2 (5%) | 6 (15%) | 40 |
| 1940 | 6 (18%) | 6 (18%) | 3 (9%) | 0 (0%) | 14 (41%) | 2 (6%) | 3 (9%) | 34 |
| 1950 | 6 (22%) | 7 (26%) | 0 (0%) | 0 (0%) | 7 (26%) | 4 (15%) | 3 (11%) | 27 |
| 1960 | 4 (17%) | 5 (22%) | 1 (4%) | 0 (0%) | 10 (44%) | 2 (9%) | 1 (4%) | 23 |
| 1970 | 2 (5%) | 16 (38%) | 3 (7%) | 1 (2%) | 16 (38%) | 4 (10%) | 0 (0%) | 42 |
| 1980 | 1 (2%) | 20 (29%) | 2 (3%) | 0 (0%) | 17 (25%) | 28 (41%) | 0 (0%) | 68 |
| 1990 | 1 (2%) | 5 (8%) | 1 (2%) | 0 (0%) | 3 (5%) | 55 (85%) | 0 (0%) | 65 |
| Total | 115 | 80 | 10 | 1 | 117 | 100 | 21 | 444 |

Source: Battalio [23, p. 73]. Copyright © 1998, Ablex Publishing Corporation.

will more readily identify elements and categories of information worthy of further study.

7. Linguistically Analyze the Data So far, data derived from the methodology described here have revealed the rhetorical contexts in which the discourse of science and technology operates. The linguistic decisions made by members of the scientific and technical professions also inform a historical study of discourse. Such an analysis will show not only the amount of specialization

in the field as evidenced by the use of technical terminology, but also writers' attitudes toward their own knowledge. A study of linguistic choices may also provide insight into social interactions both within the scientific and technical community and with outside forces.

The Linguistic Database: Linguistic analysis is time-consuming because it involves word-, phrase-, and sentence-level analyses. Consequently, depending on the

size of the database of articles, a subsample taken from the entire data set may be necessary. Here again, in order to show linguistic change over time, the researcher would use a systematic subsample, as explained in step 3.

For instance, one article each from issues one and three of the journal at 5-year intervals from 1890 to 1990 may be selected, resulting in a 40-article database. The particular article chosen will depend upon the purpose of the historical study. For example, if

Fig. 2. Plotting historical data. (Source: [51, p. 10]. Copyright©1991, IEEE.)

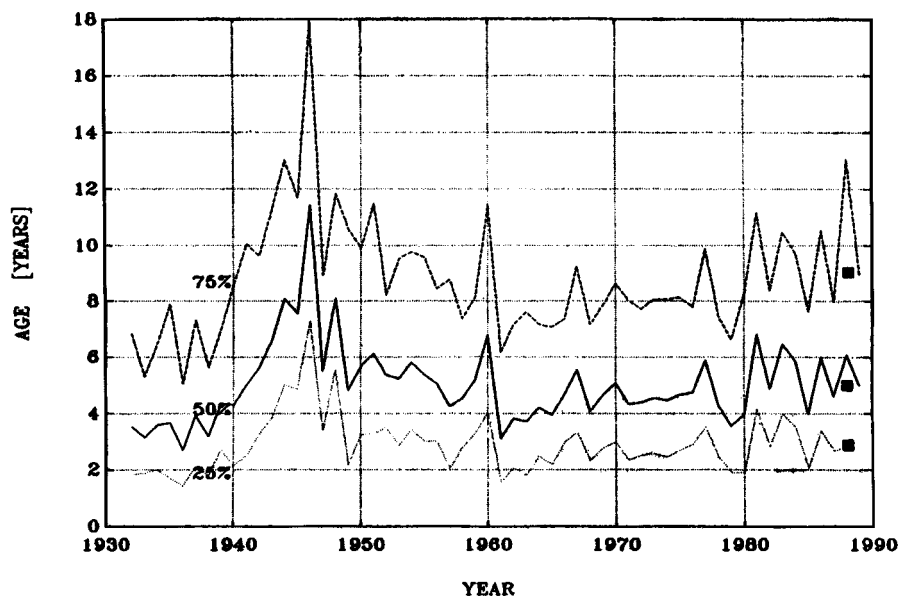
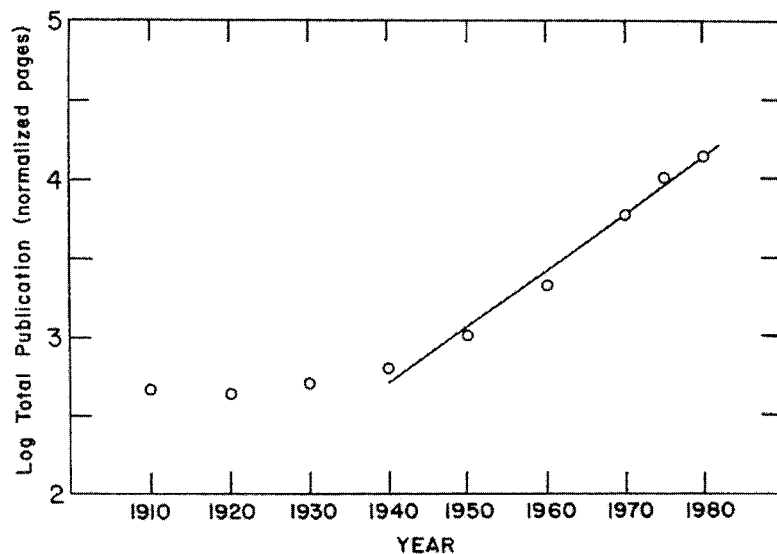


Fig. 3. Example of descriptive statistics from historical data. (Source: Abt [21, p. 271]. Copyright©1981, Astronomical Society of the Pacific.)



the field's most pressing topics are under study and if the researcher has determined by comparing the progress of research in the field over time with the journal's table of contents that article placement is indicative of a topic's priority, then the first article may be chosen. Where article placement is not a priority, the choice of article may be randomized. Since linguistic variation is also of importance, the representative sample should contain as many different authors as possible in order to sample a broad cross section of writing styles and to reduce the influence of a particular author's style of writing. Thus, during the selection process, if an author already part of the database is encountered, the researcher should not include that article. If the entire six months of articles are skipped (that is, all articles in either issues 1 and 2 or issues 3 and 4), then the first article of the six-month period (that is, the first article of issue 1 or 3) is selected. In the case of coauthored articles, selection may be based on the first author's name. This method ensures that the database contains a representative cross section of authors and subject matter, equally spread out over the years under study.

Once articles are chosen, the researcher must determine the passages to be analyzed linguistically. A useful method for selecting multiword passages for linguistic analysis is described by Bazerman, who used 1000-word passages in his analysis of physicists' discourse:

From each of the articles three to five passages for analysis were chosen, representing whichever of the following sections of the argument were present: introduction, theory, experimental, results, discussion/conclusion. The passages began at the beginning of each of the sections and ended either at the first sentence break after two hundred words were reached, or at the end of the

section if it was under two hundred words in length. [13, p. 167]

If there are no headings in the text, the first 1000 words may be selected.

Next, the researcher should consider the data gathered in previous steps and, if necessary, refine the research questions formulated at the beginning of the study. For instance, data plots from steps 5 and 6 may suggest historical implications: that sentence lengths and nominalizations will increase over time; that metadiscourse will increase in variety; that sentences will become more dense, while clauses will increase in complexity emphasizing cause-effect relationships; and/or that argumentation will be more empirically based, focusing on the use of citation, induction, and assertions of reliability.

Linguists have identified many features that may be used in describing the discourse of science and technology. Lakoff and Johnson [53] have shown the importance of the major metaphors, tropes, and figures of speech in a writer's word choice. These language features identify writers' perspectives about the world, society, and culture. Other linguistic choices, such as use of technical terminology and variations in sentence complexity, indicate a writer's audience awareness and extent of knowledge. Chafe [54] has shown that the use of evidentials reveals writers' attitudes and certainties about information, that is, expressions indicative of reliability of evidence, belief, induction, hearsay, citations, deduction, hedging, sensory evidence and expectations. Myers [22, p. 72-73] and Fahnestock [55, p. 21-30] suggest other approaches for evaluating writers' attitudes.

The Linguistic Analysis: Once the linguistic features are selected, two or more raters code into the text

for analysis and future reference each feature encountered in the sample passage. The data are then marked on a coding sheet, which is later used to count occurrences of each linguistic feature. For instance, a linguistic analysis may be made of diction choices, sentence complexity, and evidentiality. As each element is found in the relevant texts, its location is marked in the text and noted on the coding form, which is later used to compile occurrences of each linguistic feature (see [23, p. 185] for an example).

After occurrences are counted and insufficient cell counts are eliminated (see step 3), the data may be grouped into periods (for instance, 10-year segments), averaged and graphed as frequency distributions. Transforming the data into time-line graphs will show positive or negative trends in the linguistic features under study. As explained in step 5, such graphs will provide a preliminary view of the linguistic changes over the time span of the journal being analyzed in the same way that Fig. 3 shows changes in the lengths of astronomers' articles. Linguistic patterns may likewise correspond to the historical changes in the discourse.

Because linguistic study necessitates relatively small databases, time lines that show decided positive or negative slopes should be the focus of the analysis, rather than a one-period irregularity that may be caused by a particular author's overuse of the linguistic feature being analyzed. These slopes most likely point to those linguistic features indicative of the most change over time and are consequently the most useful for distinguishing the discourse of science and technology. In particular, the data associated with these slopes are the most useful for inferential statistical analysis (see step 8).

Two trends will be of particular interest. First, linguistic features

may be generally stable over time; values early in the study period are similar to those late without any large positive or negative slopes. Of course, some fluctuations in the data are likely, due to both the limited database and variation in individual writing styles. Second, linguistic features may show definite changes over time. Once these patterns are identified, the researcher now has justification for conducting further, more extensive research using more articles and larger linguistic samples over the more limited time period indicated by the plotted time lines. The resulting data may show even larger changes and steeper slopes.

Regardless of the methodology, linguistic analysis of individual articles creates the possibility of documenting only individual writers' styles because an individual writer's influence cannot be completely separated from any piece of discourse. There will always be some variance in the empirical results of such a study. Individual writers make decisions about their writing based upon past experience, the limitations imposed by the discourse, and the audience itself. The linguistic data obtained from individual authors' styles is best viewed as an indication of and generalization about the type of writing style typical of the discourse of a particular field.

8. Implement Categorical

Data Analysis The first seven steps above have shown that a researcher can obtain useful information by descriptive statistics. As Fig. 1 indicates, step 8 may be considered an optional one, depending upon the researcher's expertise in inferential statistical analysis. For an introductory overview of quantitative analysis, see Goubil-Gambrell [25] and Spyridakis et al. [56]. MacNealy [29], Lauer and Asher [26], and Spyridakis [30] provide more comprehensive descriptions.

Here I will describe the use of content analysis and commercial statistical packages. However, this overview does not pretend to substitute for a formal study of statistical analysis. On the contrary, academic study of statistics provided by many universities is a prerequisite to full use of statistical analysis. Campbell [3] offers an analysis of the strengths and weaknesses of current research-methods courses in programs in business and technical communication.

Still, personal computer software programs, such as PC SAS and advanced spreadsheet software programs, are making the job, at least of obtaining the calculations, much easier. On the other hand, those without statistical expertise may wish to work in collaboration with a colleague in the math or statistics department.

Content Analysis: Historical researchers, whose subjects and contexts may no longer be directly accessible for qualitative study, may still simulate such contact through content analysis, a means of obtaining information by indirect methods. Krippendorf [37] provides the following examples. A linguist may focus on the words and expressions found in documents. A sociologist may derive the meaning of an act by placing it in its social context. A communication researcher may document influence on audience, cultural distinctions, and institutional constraints by interpreting a message "in relation to a sender's intentions, to a receiver's cognitive or behavioral effects, to the institutions within which it is exchanged, or to the culture within which it plays a role" [37, p. 24]. Psychologists have used content analysis to determine issues such as "writer's cognition, attitudes, personality, [and] psychopathology" [37, p. 25] by analyzing personal documents.

Garraty [57] justifies and describes the use of content analysis in

historical research. Krippendorf also addresses the importance of the methodology to historical researchers:

Historians would surely love to interview Caesar, Nietzsche, or African slaves entering colonial America, if this were possible. But because letters, books, artifacts, and records kept by others do not anticipate the historian's questions, answers have to be found by indirect methods.... [H]istorians have always demanded that documents be placed in their appropriate historical context. Once this context is delineated, gaps in relevant detail are filled by drawing inferences from numerous bits of information [37, p. 24].

Hagge, who justifies the use of content analysis for making generalizations about written texts [58, pp. 408-420], briefly describes the methodology:

A researcher (a) chooses a representative sample of data to be analyzed; (b) determines, after becoming familiar with the general range of data, a method for classifying the data into mutually exclusive categories; (c) codes the data, tests the coding scheme, makes adjustments, and recodes; and (d) makes inferences about the overall content categories into which the data falls based on the previous three steps [58, p. 420].

In brief, texts may be classified by paragraph, word, word sense, sentence, or theme in accordance with results obtained in the previous steps. Texts may then be coded either by hand or by a computer programmed to implement various standardized classification themes. Krippendorf [37] provides examples that describe hand-coded methodologies. Weber [38] describes the use of computers to produce content analyses of texts. It should be noted that content

analysis should be distinguished from interpretive methodologies such as contextual or rhetorical analyses. In content analysis, the mere frequency of occurrences of words and their meanings has interpretive value. Whether hand or computer coded, counts or lists of relevant words or concepts are compiled as word frequencies, concordances, or keyword-in-context, the latter being a list of excerpts containing the coded word or some specified criteria. Numerical results may then be analyzed quantitatively.

Statistical Packages: In addition, analyses of categorical data may be made using the SAS statistical package. Other packages, such as SPSS, can perform comparable calculations. Three procedures in SAS are of use in analyzing categorical data: PROC FREQ, PROC CATMOD, and PROC LOGISTIC. PROC FREQ provides a number of measures of association, similar to correlation coefficients in continuous data analysis, that show if there are significant relationships between factors [59]. Table III shows the variety of statistical measures available from PROC FREQ. The important point here is that these measures are functions of chi-square (χ^2), the traditional statistic for count data. Statisticians have documented numerous problems with the chi-square itself, including its oversensitivity to table and sample size, scale dependence, and its inability to identify an association or its direction. The various measures in Table III offset in varying degrees the problems with the standard chi-square.

In particular, the likelihood ratio (G^2) has become popular because of its statistical properties. As a result, it is especially applicable for the log-linear procedures described below since the statistic may be used like the sum of squares in ANOVA and regression models.

PROC CATMOD provides three features of use in analyzing

categorical data. Loglinear and logit models can perform an ANOVA-like analysis of two-, three-, and four-factor tables showing the relationships among sets of data. Logistic regression models can perform regression-like analyses of multiple factors that show how strong relationships are among factors and whether these relationships are positive or negative [60]. In addition, PROC LOGISTIC also performs regression analyses of categorical data.

Thus, inferential statistical analyses such as these add double value to historical studies. First, associations among data may be revealed in a greater variety of ways. The more ways the researcher has of looking at associations among data, the greater the chance of finding significant associations. Second, though descriptive statistics may show trends among data,

inferential statistics estimates the significance of those trends and suggests the relationships among the historical, linguistic, and rhetorical components of the discourse, thereby enhancing the researcher's chances of identifying the key features of the profession's discourse.

9. Develop Schemas of Discourse At this point in the process, the researcher has identified and analyzed discourse variables and factors, along with having described rhetorical and linguistic components of the discourse. The overviews and continuums established by the researcher have put these elements into historical perspective. The researcher may now ask if it is possible to construct a diagrammatic representation, or discourse schema, that shows pictographically the interrelationship among all of the concepts discovered thus far. The mere act of creating such representations may clarify and put into perspective the many diverse findings from the historical study. Drawing a picture that synthesizes the major conclusions gleaned from the study forces the researcher to juxtapose, interconnect, and reconsider visually the interpretations that have been developing internally, perhaps subconsciously.

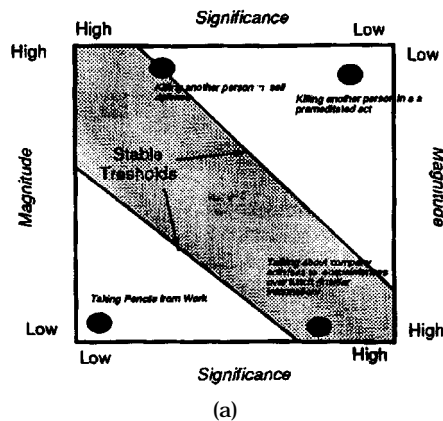
Authors often present theoretical concepts diagrammatically. For instance, the theoretical area called the Ethical Dilemma Domain is depicted as the shaded area shown in Fig. 4(a) [61]. Fig. 4(b) shows the minimalist theory of mind schema reinforcement [62]. Fig. 4(c) is the discourse schema representing the progression from natural history to experimental science ornithological discourse, shown by the horizontal arrow [23]. Since the last example is taken from a study that follows the methodology described here, a further explanation of this schema may be instructive in showing the usefulness of this step in

TABLE III
LIST OF STATISTICAL MEASURES
OF ASSOCIATION

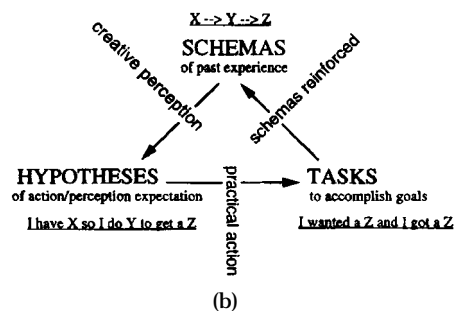
| |
|--------------------------------------|
| ANOVA (Row Mean Scores Differ) |
| Cramer's V |
| Cross-Product Ratio |
| Equality of Proportions |
| General Association |
| Goodman & Kruskal's Gamma |
| Kendall's τ_α |
| Kendall's τ_β |
| Kruskal's λ |
| Likelihood Ratio Chi-Square |
| Mantel-Haenszel's Chi-Square |
| Pearson's Chi-Square |
| Pearson's Coefficient of Contingency |
| Pearson's r |
| Phi Coefficient |
| Relative Risk |
| Somers' D |
| Spearman's ρ |
| Stuart's τ_c |
| Uncertainty Coefficient |
| Yule's Q |

the methodology. The shaded area represents discourse in transition from 1900 until 1970, the latter date approximating professionalization. Variation in the amount of shading along the Continuum indicates the level of discourse disparity. The jagged lines between 1940 and 1970 symbolize the points of statistical significance where social, historical, linguistic, and scientific forces were reacting to the approaching professionalization. Thus, the schema is, in effect, a summary of the study's findings. Such visualizations move numerical data from the abstract to the concrete plane,

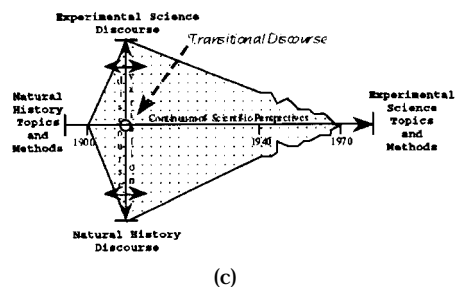
Fig. 4. Sample discourse schema. (a) Sturges [61, p. 46]. Copyright © 1992, IEEE. (b) Manning [62, p. 201], Copyright ©1998, IEEE. (c) Battalio [23, p. 163]. Copyright © 1998, Ablex Publishing Corporation.



(a)



(b)



(c)

thereby preventing the individual steps in the methodology from compartmentalizing the results.

10. Implement Qualitative Analysis Though the methodology described in the first nine steps may have uncovered valuable historical data, it is even more valuable in that it has narrowed a large textual record into more specific, manageable factors of interest—perhaps the journal articles for a decade or year, rather than a century; a set of articles by a particular author or authors, rather than a database of 200 authors; or a set of articles surrounding a particular technical or scientific event or discovery, rather than a hundred years of history.

Now, if analysis reveals that the data warrant further analysis, textual data may undergo more extensive linguistic and rhetorical analyses. Personages and events may be the subjects of qualitative analyses, all of which may be made because the historical period and textual record are more manageable. For instance, my analysis of numerous factors in ornithological discourse found a significant relationship between the professionalization of ornithology during the mid-20th century and linguistic features like the number and frequency of subject-verb interrupters, markers for reliability of evidence, technical terminology, and high-density graphics [23]. A more extensive linguistic analysis of a larger body of journal articles between 1950 and 1970 may uncover further linguistic features or clarify the ones already discovered. Qualitative studies of the ornithological researchers themselves and the significant ornithological, scientific, and socio-cultural events during these two decades may shed further light into ornithology's professionalization. These more concentrated studies are possible because the methodology described above captured and analyzed a

hundred years of data and provided results that point to those 20 years during the mid-20th century as the turning point in American ornithological discourse.

Though more quantitative study may be made of this narrower body of data, I offer an overview of the kinds of qualitative research that may now be conducted using the data acquired from the previous steps in this methodology. For those unfamiliar with qualitative research methodology, Goubil-Gambrell [25] and MacNealy [29] give an overview of the process. Though Campbell [63] focuses on conducting qualitative research in the workplace, she provides an excellent methodological overview.

I suggest here several generally accepted qualitative methodologies (surveys, case studies, and ethnographies) that would be useful in historical studies of technical and scientific discourse and I offer a number of other approaches.

Though survey methodology may not be appropriate for historical materials over long time periods since a representative sample of respondents may not be available, those studies whose methodology has led researchers to focus on contemporary periods may find surveys valuable. For instance, if a researcher finds the period between 1960 and 1980 significant, surveys are appropriate since many potential respondents are still alive. Surveys are appropriate for qualitative analysis since they may be designed to elicit personal responses in addition to numerical data. Plumb and Spyridakis [64] describe the major steps in preparing, administering, and evaluating surveys. Anderson [65], Lauer and Asher [26], and MacNealy [29] offer more comprehensive overviews.

If historical data show that the influence of particular individuals in a technical or

scientific field warrants further study, case studies of these individuals may be undertaken. Such studies may uncover, for instance, personal beliefs and motivations, professional activities, or interpersonal relationships that may assist in explaining the individual's influence and thus shed light on the discourse itself. MacNealy [29] and Lauer and Asher [26] describe this methodology.

Because context is essential in a historical study, ethnographic studies complement historical research since they analyze and describe the relationship of individuals to their environments. See Doheny-Farina and Odell [66] and MacNealy [29] for a description of this research methodology.

Other scholars have suggested nontraditional qualitative approaches. For instance, Blyler [34] suggests the use of critical research focusing on social action whereby the relationship between researchers, individuals in the study, and research goals are redefined so as to enable empowerment and emancipation. Longo [32] proposes the use of cultural studies research to extend the idea of culture beyond institutional boundaries. Researchers may then analyze how texts function in our culture's knowledge/power system, thus enlightening multicultural, gender, ethical, and other cultural issues.

Such qualitative research may be a springboard to other types of research. It may uncover data that can be in turn analyzed quantitatively, thereby giving rise to a recursive methodology. In addition, both quantitative and qualitative analyses have led to rhetorical analyses, which, as noted by Moran, define the history of scientific and technical writing more clearly than "closely related areas such as literary criticism, the history of science, and technology, the philosophy of science and the history of ideas" [6, p. 25].

Qualitative methodologies often strive to obtain an in-depth understanding of the beliefs, motivations, and attitudes of individuals and of their interactions with their associates and surroundings. Because historical researchers often study individuals who lived in distant historical periods, researchers may be unable to become participant-observers in such circumstances or to conduct opinion surveys, since the environment under study no longer exists or potential respondents are deceased. However, knowledge of environments, subjects, and their interaction may be obtained through content analysis, as described in step 8 above. Though Hagge contrasts content analysis with qualitative studies such as ethnographic and discourse studies [58, pp. 417–418], these methodologies also complement each other. That is, after the researcher accumulates personal written artifacts, such as letters, book reviews, and short commentaries, or written artifacts surrounding a particular historical period or event of interest, a content analysis may then be made of these texts in order to determine such things as the mental attitude of the writer; the influence of the document on historical events; and the evolution of theories, opinions and ideas.

Though content analysis is often conducted quantitatively, for instance, by obtaining frequency occurrences of vocabulary or grammatical usage, it may also be used qualitatively (see George [67] and Thompson [28, p. 51]). In particular, Thompson [28] shows the value of content analysis in obtaining functional, rhetorical, ideological, intercultural, and feminist perspectives.

Another advantage of the 10-step method proposed here is that by ending with qualitative analysis, the research remains open-ended. Unlike quantitative

analysis, which seeks specific answers from the numerical data, qualitative analysis, through its more expansive approaches and data collection, often finds new questions to be answered. As a result, such a study may end by suggesting the next avenue for research, thereby satisfying a major characteristic of well-designed research studies—that they advance future scholarly inquiry.

CONCLUSION

The 10-step process outlined here is one method for systematically collecting historical information from a study of a large body of scientific and technical publications. I offer it as one framework to assist in establishing technical communication as a legitimate academic discipline. Most importantly, it is a multifaceted means for analyzing a large body of historical data: it may be embedded within a qualitative study that encounters a large body of data in need of such analysis; it may be implemented as its own quantitative study; or it may be used in conjunction with other methodologies such as critical research.

To summarize the procedure briefly, the analysis begins by first determining the focus of the historical study. The researcher then searches for overviews of the field under study. Aided by this knowledge, the researcher uses systematic sampling to select a database from the entire body of work. After this database is read, its data are placed in a historical data matrix. These data are then plotted in order to search for historical patterns or to refine the data further. Next, changes in the discourse over time are identified, and linguistic analysis is implemented. Categorical data analysis may then be used to gain further insight into the implications of the data that have been collected. Discourse schemas

are developed to put the findings in perspective. By this time, the study should have revealed both historical patterns in the discourse and time periods critical in the evolution of the professional field under study. Finally, depending upon the research question under study, conclusions may be drawn or further research may ensue.

The methodology has implications for future research in that it may be used to enhance a number of previously published historical studies that employed the analysis of journal articles. For instance, the study by Allison [19] of the inability of experimental parapsychology to enter into mainstream science is based largely on member surveys. Since experimental parapsychology's methodology and research topics are two major obstacles in the field's inclusion into mainstream science, a further study might survey the evolution of these two obstacles over the life of parapsychology's two journals using the methodology described here. The study of the professionalization of American sociology by Kuklick [20] includes the use of sociology journals to document a specific topic, the continuing search for identity and status. If the full range of research topics were now sampled and analyzed as described

above, the almost 100 years of publications should hold a wealth of information into the evolution of the field. The study of British geology by O'Connor and Meadows [18] makes general assertions about the increasing specialization and exclusivity of the field by focusing on a number of amateur and professional, as well as various provincial, geological publications. A more specific analysis of a large body of journal articles, as described here, from the two major British geological societies might pinpoint a more specific time period for the break, which could then be studied in greater depth. Finally, Brasseur and Thompson [68], arguing for further historical analysis of technical communication, demonstrate gendered ideology in three sets of illustrations from two 17th-century medical manuals. This study might be expanded to an analysis of gendered ideology over the life of medical manuals by following the methodology described here. That is, a random sample of 17th-century medical manuals would be selected for analysis. From these manuals, illustrations deemed appropriate for an analysis of gender ideology would then be identified, and a random sample of these illustrations chosen for analysis. These are just four of the

many historical studies of journal articles that historical researchers may target for further research using the methodological approach just presented.

Latour [69] provides researchers studying the evolution of science a means for studying the origins and development of scientific thought, community, and discourse. Essentially, he argues for the need to move historically back to the origin of controversies—to the making of knowledge—rather than to investigate ready-made science, that is, science whose theories and precepts have already been black-boxed into certified, accepted "certainties." In other words, to understand who you are, you must know where you have come from. By researching the historical origins and evolution of the technical and scientific professions, scholars have the opportunity to describe more fully the evolution of these disciplines. The framework proposed here should assist in this discovery.

ACKNOWLEDGMENT

I thank the anonymous reviewers for their comments during the review process. I also thank Editor K. S. Campbell for her assistance during the revision of this article.

REFERENCES

- [1] L. K. Grove and D. E. Zimmerman, "Introduction: Bringing communication science to technical communication—Advancing the profession," *IEEE Trans. Prof. Commun.*, vol. 40, no. 3, pp. 157–167, 1997.
- [2] G. J. Savage, "The process and prospects for professionalizing technical communication," *J. Tech. Writing and Commun.*, vol. 29, no. 4, pp. 355–381, 1999.
- [3] K. S. Campbell, "Research methods course work for students specializing in business and technical communication," *J. Bus. Tech. Commun.*, vol. 14, no. 2, pp. 223–241, 2000.
- [4] E. Tebeaux and M. J. Killingsworth, "Expanding and redirecting historical research in technical writing: In search of our past," *Tech. Commun. Quart.*, vol. 1, no. 2, pp. 5–32, 1992.
- [5] J. J. Connor, "Medical text and historical context: Research issues and methods in history and technical communication," *J. Tech. Writing and Commun.*, vol. 23, no. 3, pp. 211–232, 1993.

- [6] M. G. Moran, "The history of technical and scientific writing," in *Research in Technical Communication: A Bibliographic Sourcebook*, M. G. Moran and D. Journet, Eds. Westport, CT: Greenwood, 1985, pp. 25-38.
- [7] T. C. Kynell and M. G. Moran, *Three Keys to the Past: The History of Technical Communication*. Stamford, CT: Ablex, 1999, vol. 7, ATTW Contemporary Series in Technical Communication.
- [8] E. Tebeaux, *The Emergence of a Tradition: Technical Writing in the English Renaissance: 1475-1640*. Amityville, NY: Baywood, 1997.
- [9] —, "Technical writing in seventeenth-century England: The flowering of a tradition," *J. Tech. Writing and Commun.*, vol. 29, no. 3, pp. 209-253, 1999.
- [10] —, "Visual texts: Format and the evolution of English accounting texts, 1100-1700," *J. Tech. Writing and Commun.*, vol. 30, no. 4, pp. 307-341, 2000.
- [11] G. Lippincott, "Experimenting at home: Writing for the nineteenth-century domestic workplace," *Tech. Commun. Quart.*, vol. 6, no. 4, pp. 365-380, 1997.
- [12] "Special Issue on Historical Contributions by Women to Technical Communication," *Tech. Commun. Quart.*, vol. 6, no. 3, pp. 249-329, 1997.
- [13] C. Bazerman, *Shaping Written Knowledge*. Madison, WI: Univ. Wisconsin Press, 1988.
- [14] C. Berkenkotter and T. N. Huckin, *Genre Knowledge in Disciplinary Communication: Cognition/Culture/Power*. Hillsdale, NJ: Lawrence Erlbaum, 1995.
- [15] J. M. Swales, *Genre Analysis: English in Academic and Research Settings*. New York: Cambridge Univ. Press, 1990.
- [16] D. Atkinson, *Scientific Discourse in Sociohistorical Context (The Philosophical Transactions of the Royal Society of London 1675-1975)*. Mahwah, NJ: Lawrence Erlbaum, 1999.
- [17] T. Orr, "Genre in computer science and computer engineering," *IEEE Trans. Prof. Commun.*, vol. 42, no. 1, pp. 32-37, 1999.
- [18] J. G. O'Connor and A. J. Meadows, "Specialization and professionalization in British geology," *Social Studies of Science*, vol. 6, pp. 77-89, 1976.
- [19] P. D. Allison, "Experimental parapsychology as a rejected science," in *On the Margins of Science: The Social Construction of Rejected Knowledge*, R. Wallis, Ed. Keele, U.K.: Univ. Keele, 1979, pp. 271-291. *Sociological Review Monograph 27*.
- [20] H. Kuklick, "Boundary maintenance in American sociology: Limitations to academic professionalization," *J. Hist. Behav. Sci.*, vol. 16, no. 3, pp. 201-219, 1980.
- [21] H. A. Abt, "Some trends in American astronomical publications," *Publ. Astronom. Soc. Pacific*, vol. 93, no. 553, pp. 269-272, June 1981.
- [22] G. Myers, *Writing Biology: Texts in the Social Construction of Scientific Knowledge*. Madison, WI: Univ. Wisconsin Press, 1990.
- [23] J. T. Battalio, *The Rhetoric of Science in the Evolution of American Ornithological Discourse*. Stamford, CT: Ablex, 1998, vol. 8, ATTW Contemporary Series in Technical Communication.
- [24] J. Detweiler and C. Peyton, "Defining occupations: A chronotopic study of narrative genres in a health discipline's emergence," *Written Commun.*, vol. 16, no. 4, pp. 412-468, 1999.
- [25] P. Goubil-Gambrell, "A practitioner's guide to research methods," *Tech. Commun.*, vol. 39, no. 4, pp. 582-591, 1992.
- [26] J. M. Lauer and J. W. Asher, *Composition Research: Empirical Designs*. New York: Oxford Univ. Press, 1988.
- [27] P. Sullivan and R. Spilka, "Qualitative research in technical communication: Issues of value, identity and use," *Tech. Commun.*, vol. 39, no. 4, pp. 592-606, 1992.
- [28] I. Thompson, "Competence and critique in technical communication: A qualitative content analysis of journal articles," *J. Bus. Tech. Commun.*, vol. 10, no. 1, pp. 48-80, 1996.
- [29] M. S. MacNealy, *Strategies for Empirical Research in Writing*. Boston, MA: Allyn & Bacon, 1999.
- [30] J. H. Spyridakis, "Conducting research in technical communication: The application of true experimental designs," *Tech. Commun.*, vol. 39, no. 4, pp. 607-624, 1992.

- [31] D. Charney, "Empiricism is not a four-letter word," *College Comp. Commun.*, vol. 47, no. 4, pp. 567-593, 1996.
- [32] B. Longo, "An approach for applying cultural study theory to technical writing research," *Tech. Commun. Quart.*, vol. 7, no. 1, pp. 53-73, 1998.
- [33] M. Markel, "Induction, social constructionism and the form of the scientific paper," *J. Tech. Writing and Commun.*, vol. 23, no. 1, pp. 7-22, 1993.
- [34] N. Blyler, "Taking a political turn: The critical perspective and research in professional communication," *Tech. Commun. Quart.*, vol. 7, no. 1, pp. 33-52, 1998.
- [35] B. Berelson, *Content Analysis in Communication Research*. Glencoe, IL: Free Press, 1952.
- [36] I. de Sola Pool, *Trends in Content Analysis*, I. de Sola Pool, Ed. Urbana, IL: Univ. Illinois Press, 1959.
- [37] K. Krippendorff, *Content Analysis: An Introduction to Its Methodology*. Beverly Hills, CA: Sage, 1980.
- [38] R. P. Weber, *Basic content analysis*, 2nd ed. Newbury Park, CA: Sage, 1990. Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-049.
- [39] D. C. Miller, *Handbook of Research Design and Social Measurement*, 5th ed. Newbury Park, CA: Sage, 1991.
- [40] L. Ott, *An Introduction to Statistical Methods and Data Analysis*, 3rd ed. Boston, MA: PWS-KENT, 1988.
- [41] L. Anthony, "Writing research article introductions in software engineering: How accurate is a standard model?," *IEEE Trans. Prof. Commun.*, vol. 42, no. 1, pp. 38-46, 1999.
- [42] B. Wahlstrom and M. Lay, "Editors' column," *Tech. Commun. Quart.*, vol. 1, no. 1, pp. 5-6, 1992.
- [43] G. F. Hayhoe, "Evolution • revolution: Changing the way our profession communicates," *Tech. Commun.*, vol. 43, no. 2, pp. 126-129, 1996.
- [44] National Research Council, *Physics in Perspective*, ser. The Interfaces. Washington, DC: Nat. Acad. Sci., 1973, vol. IIB.
- [45] C. F. Batchelder, "An account of the Nuttall Ornithological Club: 1873 to 1919," in *Contributions to the History of American Ornithology. Natural Sciences in America*, K. B. Sterling, Ed. New York: Arno, 1974.
- [46] F. Chapman and T. S. Palmer, *Fifty Years' Progress in American Ornithology, 1883-1933*. Lancaster, PA: Amer. Ornithologists' Union, 1933. In *Contributions to the History of American Ornithology, Natural Sciences in America*, K. B. Sterling, Ed. New York: Arno, 1974.
- [47] M. J. Killingsworth and J. S. Palmer, *Ecospoken: Rhetoric and Environmental Politics in America*. Carbondale, IL: Southern Illinois Univ. Press, 1992.
- [48] W. G. Cochran, *Sampling Techniques*, 3rd ed. New York: Wiley, 1977.
- [49] A. Agresti, *Categorical Data Analysis*. New York: Wiley, 1990, Wiley Series in Probability and Mathematical Statistics, Applied Probability and Statistics.
- [50] M. V. Battle, G. S. M. Cowan Jr., and E. A. Rakow, "Statistical power: Planning for it, estimating it post hoc, and reporting it," *IEEE Trans. Prof. Commun.*, vol. 43, no. 4, pp. 345-354, 2000.
- [51] E. Matricciani, "The probability distribution of the age of references in engineering papers," *IEEE Trans. Prof. Commun.*, vol. 34, no. 1, pp. 7-12, 1991.
- [52] D. Paul, "In citing chaos: A study of the rhetorical use of citations," *J. Bus. Tech. Commun.*, vol. 14, no. 2, pp. 185-222, 2000.
- [53] G. Lakoff and M. Johnson, *Metaphors We Live By*. Chicago, IL: Univ. Chicago Press, 1980.
- [54] W. Chafe, "Evidentiality in English conversation and academic writing," in *Evidentiality: The Linguistic Coding of Epistemology*, W. Chafe and J. Nichols, Eds. Norwood, NJ: Ablex, 1986.
- [55] J. Fahnestock, "Accommodating science: The rhetorical life of scientific facts," in *The Literature of Science: Perspectives on Popular Scientific Writing*, M. W. McRae, Ed. Athens, GA: Univ. Georgia Press, 1993, pp. 17-36.
- [56] J. H. Spyridakis, M. J. Wenger, and S. H. Andrew, "The technical communicator's guide to understanding statistics and research design," *J. Tech. Writing and Commun.*, vol. 21, no. 3, pp. 207-219, 1991.

- [57] J. A. Garraty, "The application of content analysis to biography and history," in *Trends in Content Analysis*, I. de Sola Pool, Ed. Urbana, IL: Univ. Illinois Press, 1959, pp. 171-187.
- [58] J. Hagge, "The value of formal conventions in disciplinary writing: An axiological analysis of professional style manuals," *J. Bus. Tech. Commun.*, vol. 8, no. 4, pp. 408-461, 1994.
- [59] SAS Institute Inc., *SAS/STAT Software: Changes and Enhancements Through Release 6.12*. Cary, NC: SAS Institute Inc., 1997.
- [60] D. W. Hosmer and S. Lemeshow, *Applied Logistic Regression*. New York: Wiley, 1989.
- [61] D. L. Sturges, "Overcoming the ethical dilemma: Communication decisions in the ethic ecosystem," *IEEE Trans. Prof. Commun.*, vol. 35, no. 1, pp. 44-50, 1992.
- [62] A. D. Manning, "Minimalism beyond computer documentation," *IEEE Trans. Prof. Commun.*, vol. 41, no. 3, pp. 200-204, 1998.
- [63] K. S. Campbell, "Collecting information: Qualitative research methods for solving workplace problems," *Tech. Commun.*, vol. 46, no. 4, pp. 532-545, 1999.
- [64] C. Plumb and J. H. Spyridakis, "Survey research in technical communication: Designing and administering questionnaires," *Tech. Commun.*, vol. 39, no. 4, pp. 625-638, 1992.
- [65] P. V. Anderson, "Survey methodology," in *Writing in Nonacademic Settings*, L. Odell and D. Goswami, Eds. New York: Guilford, 1985, pp. 453-501.
- [66] S. Doheny-Farina and L. Odell, "Ethnographic research on writing: Assumptions and methodology," in *Writing in Nonacademic Settings*, L. Odell and D. Goswami, Eds. New York: Guilford, 1985, pp. 503-535.
- [67] A. L. George, "Quantitative and qualitative approaches to content analysis," in *Trends in Content Analysis*, I. de Sola Pool, Ed. Urbana, IL: Univ. Illinois Press, 1959, pp. 7-32.
- [68] L. E. Brasseur and T. L. Thompson, "Gendered ideologies: Cultural and social contexts for illustrated medical manuals in Renaissance England," *IEEE Trans. Prof. Commun.*, vol. 38, no. 4, pp. 204-215, 1995.
- [69] B. Latour, *Science in Action: How to Follow Scientists and Engineers Through Society*. Cambridge, MA: Harvard Univ. Press, 1987.

John T. Battalio is an associate professor who teaches undergraduate and graduate courses in the Technical Communication Program at Boise State University. He has researched the historical and linguistic evolution of the writing of ornithologists and the interconnections between professional and popular writing in the sciences. The results of his research have recently appeared in *The Rhetoric of Science in the Evolution of American Ornithological Discourse* published by Ablex Press. He has also edited *Essays in the Study of Scientific Discourse: Methods, Practice and Pedagogy* and has published articles in *Technical Communication Quarterly* and *The Journal of Technical Writing and Communication*.