

Quantitative Genre Analysis of Undergraduate Theses: Uncovering Different Ways of Writing and Thinking in Science Disciplines

JASON E. DOWD, ROBERT J. THOMPSON, JR., AND
JULIE A. REYNOLDS

Introduction

One of the challenges for writing in the disciplines (WID) programs and practitioners has been to replace the commonly held view in academia that disciplinary conceptual knowledge is a specialized skill but writing is a generalized skill. This view stems from the understanding of disciplines primarily as domains of specialized content knowledge. Disciplinary knowledge, however, includes both conceptual content knowledge (i.e., *knowing that*) and procedural knowledge (i.e., *knowing how*). Learning in a discipline is sometimes understood as acquiring content knowledge, whereas the focus of WID “tends to be on procedural knowledge, writing as a way of knowing in a discipline” (Carter, 2007, p. 387). More specifically, “WID developed as a response to the recognition that different disciplines are characterized by distinct ways of writing and knowing. Thus, a specialized conception of disciplinary knowledge is integrated with a specialized conception of writing” (Carter, 2007, p. 387). Carter has drawn on the idea of disciplinary ways of doing (Herrington, 1981; Russell, 1997) as “a link between ways of writing and ways of knowing in the disciplines” (2007, p. 387).

A second challenge for WID programs has been to characterize and teach salient differences in disciplinary ways of knowing, doing, and writing. Ways of knowing and doing include epistemic beliefs, methods of inquiry, and processes for making meaning of findings. Experts across disciplines utilize different paradigms, or “system[s] of beliefs and practices that guide a field” (Morgan, 2007, p. 49), by specifying the following: appropriate problems for study and research questions that are considered important and solvable; types of questions that are asked; methods that are employed and considered valid; and ways of making meaning (Kalman, 2009). More specifically, experts across disciplines differ in epistemic beliefs about the nature of knowledge and the truth criteria employed for the justification of claims, as well as the methods and cognitive processes involved in the construction of knowledge. At the broadest level, these disciplinary differences are evident in comparing the humanities

and sciences (Phillips & Burbules, 2000). The humanities focus on human thought and action as text, employ metaphor and the cognitive processes of analogy and intuition, and rely on the method of hermeneutics (i.e., the interpretation of meaning). The sciences, in contrast, focus on the natural and social worlds, employ rationalism and empiricism and the cognitive processes of deductive and inductive reasoning, and rely on objective evidence that is publicly available for inspection and replicable by the relevant professional community (Phillips & Burbules, 2000).

While broad differences between the humanities and the sciences are most readily apparent, there are also differences in ways of knowing and writing between and within the physical/life and social/behavioral science disciplines that employ scientific methods. Some science disciplines pose “what” questions, while others pose “how” questions, to explain the mechanisms by which certain processes take place. For example, the concept of “mechanism” is particularly important for understanding the biological sciences (Machamer, Darden, & Craver, 2000). Explanation relies upon isolating causal factors, and science disciplines differ in their reliance on experimental or statistical control. Some disciplines rely more heavily on deductive reasoning and engage in “hypothetico-deductive” theory building (e.g., theoretical/mathematical physics), whereas other disciplines rely on inductive reasoning and model building (e.g., biochemistry or economics) (Thagard, 2006). Even within a particular science discipline, multiple approaches to reasoning may be appropriate in different contexts. This has led to the recognition that there is no single scientific epistemology (Rudolph, 2000; Russ, 2014). Students must learn and become enculturated into the ways of knowing, doing, and writing of specific scientific disciplinary communities of discourse.

One method for characterizing differences in disciplinary ways of knowing, doing, and writing is in terms of *genre*, which refers to patterns in social action of language in response to recurring rhetorical situations (Carter, 2007). Examples of genres of writing in science disciplines include laboratory reports, design/research application papers, and literature reviews of previous research. Kovac has argued that scientific research itself can be thought of as a set of rhetorical tasks: “What needs to be explained? What constitutes an explanation? How does an explanation constrain what counts as evidence” (2003, p. 237). Scientific writing, then, is a form of rhetoric in that scientists construct arguments—first to persuade themselves, and then others—that an interpretation of findings or an explanation is valid and interesting. Kovac also noted that the form that scientific writing takes reflects differences in approach: “Experimental articles are usually written so as to suggest that science is inductive, whereas theoretical articles are written to suggest that all theory is deductive” (2003, p. 236). Disciplines employ different rhetorical devices, such as metaphors. Chemistry, for example, may be considered a “metaphor-rich science” since

chemists develop models or metaphors to describe the behavior of electrons, atoms, and molecules (Kovac, 2003).

Recently, Wolfe and colleagues (2015) developed a method, Comparative Genre Analysis (CGA), to illuminate both rhetorical differences among disciplines and the manner in which particular rhetorical conventions support disciplinary ways of knowing and core intellectual values. CGA can also serve as a pedagogical strategy by making visible to students the links between genre conventions and discipline-specific values and ways of knowing. In particular, Wolfe and colleagues (2015) describe three dimensions of rhetorical analyses of different academic genres that reflect these links. One dimension relates to the methods and conceptual lenses that are used in a discipline to identify patterns and formulate meaningful interpretations. In this context, lenses are concepts, theories, or hypotheses that are used to organize and interpret the phenomena being studied. Although goals may vary across disciplines, common elements include applying lenses to interpret primary material, evaluating the relationship between study results and pre-existing concepts or hypotheses, and reflecting on the original concept. A second dimension relates to macrostructures that are used to arrange arguments. A macrostructure is an organizational pattern that provides readers with a frame of reference that helps them recall information and make sense of the text. For example, within laboratory and experimental reports, the macrostructures of introduction, methods, results, and discussion are the common norms across disciplines that convey new information in a predictable format. Within literature reviews, multiple different macrostructures are evident. For example, Wolfe and colleagues (2015) found that the macrostructure of thesis-statement-first (i.e., beginning a section with one's thesis statement) is characteristic in psychology, whereas thesis-statement-last is more typical in biology. A third dimension relates to naming and citation conventions that differ across disciplines within the same genres, including the use/avoidance of passive voice, when and how to cite other authors, and whether to use direct quotations. These conventions reflect different disciplinary values and assumptions about research and authorship, such as whether to focus on ideas and findings or theorists and researchers. Disciplines differ in the extent to which they privilege individuality and particularity of knowledge, as well as how they handle controversy (i.e., whether to cite other authors by name to refute claims or focus on knowledge claims and alternative explanations) (Wolfe et al., 2015).

The use of genre analysis as a method of elucidating and teaching disciplinary-specific values, ways of knowing, and ways of writing is encouraging. The qualitative categories and descriptors discussed above provide a robust framework for discussion of differences. The aim of the current study is to contribute to this use of genre analysis, specifically through the use of quantitative methods for analyzing emergent differences across disciplines. Our work is focused on assessing and enhancing students'

scientific reasoning and writing within the genre of the undergraduate thesis. In this study, we investigate disciplinary rhetorical differences and discuss our findings in relation to the three rhetorical dimensions of the CGA—conceptual lens, macrostructures to construct arguments, and naming and citation conventions—to further elucidate the general and discipline-specific nature of scientific reasoning and writing. We will discuss potential implications of our findings for educational practice, and we will highlight future studies.

Science Reasoning in Thesis Writing Across Disciplines

At Duke University, we have used the genre of the undergraduate honors thesis as the rhetorical context in which to study and improve students' scientific reasoning and writing. We view the process of writing an undergraduate honors thesis as a form of professional development in the sciences (i.e., a way of engaging students in the practices of a community of discourse). We have found that structured courses designed to scaffold the thesis-writing process and promote metacognition can improve writing and critical thinking skills in biology, chemistry, and economics (Dowd, Connolly, Thompson, Jr., & Reynolds, 2015; Dowd, Roy, Thompson, Jr., & Reynolds, 2015; Reynolds & Thompson, 2011). To enhance these structured courses, we developed the Thesis Assessment Protocol (TAP) as a tool for facilitating communication (the full text of the TAP, as implemented in Biology, is published in Reynolds & Thompson [2011] supplemental materials). The TAP systematically guides students and faculty through a “draft-feedback-revision” writing process, modeled after professional scientific peer review processes. The TAP includes activities and worksheets that allow students to engage in critical peer review, and provides detailed descriptions of the questions (i.e., dimensions) upon which such review should focus; these descriptions are presented as rubrics. Nine rubric dimensions focus on communication to the broader scientific community, and four rubric dimensions focus on evaluation by topic-specific experts (Table 1). These rubrics provide criteria by which the thesis may be assessed, and therefore allow the TAP to be used as an assessment tool as well as a teaching resource (Reynolds, Smith, Moskovitz, & Sayle, 2009).

Table 1. Theses assessment protocol dimensions.

Scientific community-focused	1	Is the writing appropriate for the target audience?
	2	Does the thesis make a compelling argument for the significance of the student's research within the context of the current literature?
	3	Does the thesis clearly articulate the student's research goals?
	4	Does the thesis skillfully interpret the results?
	5	Is there a compelling discussion of the implications of findings?
	6	Is the thesis clearly organized?
	7	Is the thesis free of writing errors?
	8	Are the citations presented consistently and professionally throughout the text and in the list of works cited?
	9	Are the tables and figures clear, effective, and informative?
Topic-specific	10	Does the thesis represent the student's significant scientific research?
	11	Is the literature review accurate and complete?
	12	Are the methods appropriate, given the student's research question?
	13	Is the data analysis appropriate, accurate, and unbiased?

In addition to facilitating communication and formative assessment during the writing of the thesis, the TAP functions as a method for assessment of the final product. In our work, we have used the TAP to quantitatively assess the relationship between thesis-writing courses (or specific interventions within the courses) and the strength of students' science reasoning in writing across different science disciplines: economics (Dowd, Connolly, et al., 2015); chemistry (Dowd, Roy, et al., 2015); and biology (Reynolds & Thompson, 2011). In our assessment of theses for research purposes (as opposed to assessment by instructors and thesis supervisors), we have focused exclusively on the nine dimensions related to reasoning and writing (Questions 1–9). The other four dimensions (Questions 10–13) are intended to be used by the student's thesis supervisor and are not included in our analyses of theses. In each case, we have found that the TAP is effective for facilitating instructors' feedback on thesis drafts and assessing students' scientific reasoning and writing within the thesis genre. Moreover, we have investigated whether meaningful constructs underlie thesis scores on the nine dimensions of the TAP that are of interest. We conducted exploratory factor analysis of students' theses in economics and chemistry and found one dominant underlying factor in each discipline; we termed the factor *scientific reasoning in writing* (Dowd, Connolly, et al., 2015; Dowd, Roy, et al., 2015). That is, each of the nine dimensions could be understood as reflecting, in different ways and to different degrees, the construct of *scientific reasoning in writing*, and this single underlying construct was evident in the theses in both chemistry and economics.

We are now well-positioned to expand prior studies *within* each discipline and make comparisons *across* disciplines. The purpose of the current study is to investigate

whether the TAP reveals meaningful similarities and differences in disciplinary ways of thinking, doing, and writing across three disciplines that are representative of the life sciences (biology), the physical sciences (chemistry), and the social sciences (economics). More specifically, the aims are to investigate 1) whether the underlying single factor derived from the nine dimensions of the TAP in economics and chemistry theses is also evident in biology theses; and 2) whether the single underlying factor is measuring the same construct (i.e., the factors have the same meaning and implications) within each discipline. In other words, do the nine dimensions of the TAP have a similar pattern of relative contribution (emphasis) to the construct of scientific reasoning in writing in each discipline, or do the patterns suggest that the underlying factor of scientific reasoning in writing is constituted differently across these three disciplines?

Methods

Study Sample

The study sample data for this current work was comprised of the theses submitted by students at Duke University that were assessed using the TAP rubric and procedures described above and reported in previous studies: 190 theses submitted to the department of biology from 2005 to 2008 (Reynolds & Thompson, 2011); 93 theses submitted to the department of chemistry from 2000 to 2013 (Dowd, Roy, et al., 2015); and 244 theses submitted to the department of economics from 2001 to 2011 (Dowd, Connolly, et al., 2015) .

Thesis Assessment

Each of the nine dimensions were scored on a scale from 1 to 5, where a rating of 1 indicates that the dimension under consideration is either missing, incomplete, or below the minimum acceptable standards for the department and a rating of 5 indicates that the dimension is excellent and the work exhibits mastery. Graduate students and postdoctoral fellows were hired to read and rate the majority of the theses, though some were also rated by faculty; theses were rated by raters with relevant disciplinary backgrounds (e.g., biology theses were rated by raters in biology-related disciplines). Each rater completed more than eight hours of training in the use of the TAP rubric, which included examination of sample writings and assessment and discussion of complete theses that were not part of the data set for calibration. Every thesis was read and independently assessed by two raters; raters then discussed their individual scores for a thesis with one another, explained their respective reasons for any differences, and formed a single *consensus* score for each dimension. The consensus is not

an average, but rather the result of discussion; it may be one rater's score, the other's score, or an entirely different value. Raters generated a consensus score for each of the nine dimensions for each thesis; individual dimensions could be summed to yield a total score.

Interrater reliability statistics for the discipline-specific subsets of these data have been reported in prior studies.¹ For the current study, we determined interrater reliability for the pooled sample of 527 theses. The Pearson correlation coefficient between raters' pre-discussion scores is 0.88 for total scores, and range from 0.63 to 0.79 for the nine distinct dimensions. Raters' post-discussion consensus scores are 100% in agreement. We note that scores on each dimension are within one point of each other in 86% of cases. Taken as a whole, these results indicate moderate to strong pre-discussion agreement and very strong post-discussion agreement between raters using the TAP rubric. Consensus scores were used in all analyses.

Analyses

We generated summary descriptive statistics (means and standard deviations) for the ratings of the nine TAP dimensions of students' theses in each discipline. Then, two new factor analyses were conducted. First, to determine whether there were any meaningful constructs underlying the relationships among the nine TAP dimensions of theses in biology, we conducted an exploratory factor analysis just as we had previously done in chemistry (Dowd, Roy, et al., 2015) and economics (Dowd, Connolly, et al., 2015). Second, we subsequently carried out confirmatory factor analyses of the theses in all three disciplines to determine whether the derived factors measure the same construct (i.e., have the same meaning and implications) across disciplines. To make this determination, we used the test of factorial invariance (χ^2) to evaluate whether the factor loadings for the nine dimensions of the TAP (i.e., the weight with which each dimension contributes to the single underlying factor) are the same (i.e., invariant) across the theses in biology, chemistry, and economics.² We first compared factor loadings for all nine dimensions across all three disciplines simultaneously, and then between each pair of disciplines. Finally, we examined the proportion of variance accounted for by the nine dimensions in the underlying factors for each discipline.

Results

In Table 2, we highlight means and standard deviations for the ratings of the nine assessed dimensions of students' theses in each discipline.³ These data reveal a range of performances across dimensions within each discipline. For example, students in chemistry demonstrated a relatively higher level of mastery on the dimension of *minimizing writing errors* than on *appropriate use of citations*. Regardless, relative level of performance on a specific dimension does not reflect the relative contribution of

that dimension to the underlying construct of scientific reasoning in writing in that discipline.⁴ Rather, underlying factors depend on the relationships among students' scores on different dimensions (stronger relationships correspond to stronger factor loadings).

Table 2. Summary statistics of the nine assessed dimensions of students' theses.^a

Dimension	Biology <i>n</i> = 190	Chemistry <i>n</i> = 93	Economics <i>n</i> = 244
Appropriate for audience	4.0 (1.0)	4.3 (0.9)	4.5 (1.1)
Formulating compelling argument	4.3 (1.0)	4.0 (1.0)	3.9 (1.3)
Articulating goals	4.3 (1.0)	4.2 (0.9)	4.1 (1.3)
Interpretation of results	4.3 (1.0)	3.7 (1.1)	4.2 (1.3)
Implications of results	4.1 (1.1)	3.5 (1.1)	3.7 (1.3)
Organization of thesis	4.3 (1.0)	4.0 (1.2)	4.3 (1.2)
Minimizing writing errors	4.5 (0.9)	4.7 (0.6)	4.7 (1.0)
Appropriate citation	3.4 (1.8)	1.5 (1.1)	3.6 (1.6)
Effective tables and figures	3.7 (1.4)	3.2 (1.1)	3.6 (1.2)

^a Mean values (and standard deviations, in parentheses) are shown.

Similar to previous findings in chemistry and economics, the exploratory factor analysis of TAP scores for biology theses yielded one dominant underlying factor that we have termed *scientific reasoning in writing* (Table 3). Thus, in each discipline, a single underlying factor has been identified. However, each of the factors is related to the nine dimensions of the TAP to varying degrees. When we compare across all three disciplines, we find that the factor loadings were significantly different across the disciplines ($p < 0.00001$). When we compare pairs of disciplines, we find no significant difference between the models for the factor loadings in biology and chemistry ($p = 0.68$); however, the comparisons between the models for biology and economics and between chemistry and economics were significantly different ($p < 0.00001$). Thus, the findings indicate that the underlying factors in biology and chemistry are not significantly different from one another, and the factor in economics is different from the other two factors.

To identify the nature of these differences in the constructs underlying the factors across the three disciplines, we examined the relationships reflected in the factor loadings of each of the nine dimensions of the TAP with the underlying factors (Table 3). Although we identify a single underlying factor in each discipline, the factors account for different amounts of the total variance, ranging from 22% in biology to 26% in chemistry to 52% in economics (Table 3). Additionally, the nine items load onto each factor to varying extents. Therefore, although a single dominant factor,

scientific reasoning in writing, was identified for the theses in each discipline, the relative contribution of dimensions that make up each of the three factors varies. As shown in Table 3, the dimensions that load most strongly in biology are *formulating compelling argument* and *articulating goals*, whereas the strongest dimensions in chemistry involve the *interpretation* and *implication* of the results. The factor loadings across the nine dimensions in economics are more uniformly strong, except for the relative weakness of the loading for *citations*. Structural attributes of the writing, such as *organization* and *minimizing writing errors*, are more strongly related to the factor in economics than to the factors in biology and chemistry.

Table 3. Factor loadings of the nine assessed dimensions of students' theses.

Dimension	Biology	Chemistry	Economics
Appropriate for audience	0.48	0.46	0.69
Formulating compelling argument	0.67	0.50	0.75
Articulating goals	0.54	0.40	0.75
Interpretation of results	0.51	0.65	0.80
Implications of results	0.50	0.81	0.70
Organization of thesis	0.41	0.55	0.80
Minimizing writing errors	0.36	0.33	0.83
Appropriate citation	0.31	0.38	0.49
Effective tables and figures	0.20	0.25	0.63
Factor eigenvalue	1.9	2.3	4.7
Variance explained	22%	26%	52%

Note. Each discipline revealed a single underlying factor that we have termed *scientific reasoning in writing*. In biology, the dimension that loads most strongly onto the underlying factor is *structuring argument*. In chemistry, *implications of results* loads most strongly onto the underlying factor. In economics, factor loadings are uniformly strong across dimensions, except for the relative weakness of *appropriate citations*.

Discussion

The findings of this study provide support for the use of TAP as a method of quantitative thesis genre analysis that reveals disciplinary differences in thinking and writing across three science disciplines—biology, chemistry, and economics—that are of relevance for WID programs. The TAP scores of students' theses in these disciplines across nine dimensions yield single underlying factors in each discipline that we have termed *scientific reasoning in writing*. However, this factor of *scientific reasoning in writing* is not expressed in the same way across all three disciplines.

There are both disciplinary-specific and general components to scientific reasoning in writing, with the disciplinary-specific components perhaps relating to disciplinary-specific aspects of epistemic beliefs and paradigms. For example, epistemic beliefs have been linked to students' use of argumentation in science (Nussbaum, Sinatra, & Poliquin, 2008). It has been argued that disciplinary-specific aspects of the nature of science (and related epistemological beliefs) should emerge from content in the disciplines instead of being imposed as general features (Van Dijk, 2014).

With the findings of the current study, we can interpret the common single underlying factor in each discipline as reflecting a general element of scientific reasoning in writing; at the same time, we can interpret the differential loadings and the relative portion of variance accounted for by the nine contributing dimensions as indicating possible disciplinary-specific elements. The findings also indicate that the general underlying factors of scientific reasoning in writing in biology and chemistry are more similar to each other than they are to the underlying factor in economics. Even within biology and chemistry, meaningful differences occur in some of the factor loadings in these disciplines. Specifically, the contribution of *formulating a compelling argument* is stronger in biology, and the contribution of *implication of results* is stronger in chemistry.⁵ We note that the variance explained by the underlying factors in biology and chemistry are relatively low (Table 3). Although there is clear evidence for one factor in each discipline, that factor does not account for a substantial amount of the variation in scores. In other words, the individual dimensions still convey substantial unique information about student exhibition of science reasoning in writing in these chemistry and biology theses.

These thesis genre analysis findings can be interpreted in light of Wolfe and colleagues' (2015) use of CGA to describe differences in disciplinary rhetorical conventions. The CGA rhetorical dimension of *conceptual lens* refers to concepts, theories, or hypotheses that are used to organize and interpret the phenomena being studied. This organizing and interpreting function is reflected in the TAP dimensions of formulating a compelling argument for significance of research, articulation of research goals, interpretation of results, implications of findings, and effective use of tables and figures. The TAP dimensions related to thesis structure (organization), effective use of tables and figures, and writing for an appropriate audience serve a similar function as the CGA dimension of macrostructure by providing a frame of reference to help readers recall information and make sense of the text. Similarly, both TAP and CGA assess the use of citations, though there are substantial differences in the assessment and related implications across both methods. Although TAP was not initially designed for genre analysis, the findings of this study indicate that it can function as a quantitative tool for comparative analysis across disciplines within the genre of the undergraduate thesis, and provide a view of disciplinary differences that complements

qualitative analyses. Furthermore, interpreting TAP findings in terms of general and disciplinary-specific rhetorical dimensions enhances the utility of the TAP within WID contexts as a tool for facilitating communication and systematically guiding students and faculty through a “draft-feedback-revision” thesis writing process.

Building on this interpretation and the numerical values presented in Table 3, within the genre of the undergraduate honors thesis, scientific reasoning in economics writing appears to be characterized by emphasis on both the lens dimension of interpretation of results and the macrostructure dimension of thesis organization to convey new information, as well as the minimization of writing errors. We have no empirical evidence to make meaning of the seemingly critical role of minimizing writing errors, though we speculate that, in economics at Duke University, there is a relatively strong focus on learning to write alongside writing-to-learn. Scientific reasoning in biology writing is characterized by an emphasis on the lens dimension of formulating a compelling argument, and scientific reasoning in chemistry writing is characterized by an emphasis on the lens dimension of implications of results. Scientific reasoning in both biology and chemistry writing are less characterized by the macrostructure dimension of tables and figures than in economics.

These characterizations of writing, in turn, relate to disciplinary thinking. We argue that, in economics, students’ research generally involves the proposition and development of a model, the analysis of data related to this model, and the interpretation of this analysis. The theoretical framework is self-contained in a way that is unique compared to natural sciences like biology and chemistry, and therefore scientific reasoning in writing leans more heavily on students’ organization, attention to audience, and effective use of tables and figures. In biology, which is a field in which theoretical frameworks differ considerably across sub-disciplines, we suggest that students are usually employing a pre-existing model that nonetheless must be presented and justified. Thus, the argument is essential, and organization and effective use of tables and figures, while still important to a well-constructed thesis, are less discriminatory or informative with regard to underlying science reasoning and evaluative thinking. In chemistry, where theoretical frameworks are more consistent and coherent across sub-disciplines, students may not even consider the models employed as anything other than objective, coherent descriptions of the natural world. In other words, the argument may be treated as evident once the facts and principles are stated. Instead, we argue that the students’ discussion of their own research, and particularly the implications of their results, is the primary component of the chemistry thesis requiring scientific reasoning and evaluativist thinking. While tables and figures are certainly important in the presentation of results, they may tend to be more or less effective in chemistry for idiosyncratic reasons (unrelated to science reasoning in writing) compared to economics. Interpretation of results is the primary goal of

the discussion section and addresses “how” and “why” questions (Robinson, Stoller, Costanza-Robinson, & Jones, 2008). This could explain the relatively large differences in the contributions of the dimensions of *implications of results* and *formulating a compelling argument* in the factors characterizing students’ scientific reasoning in chemistry and biology. Therefore, we see plausible explanations for both general and disciplinary-specific components when we consider thesis writing within these disciplines. If we are looking for evaluativist thinking, we would expect to find it in different places in different disciplines.

There are some limitations to the data that we must acknowledge as we make the arguments presented here. We’ve analyzed hundreds of theses collected over many years, but the representation from different years is not evenly distributed within our sample. Although we have no hypothesis to explain why there would be differences in how students exhibit scientific reasoning over time, such differences could manifest as disciplinary differences in this analysis. In comparing Tables 2 and 3, one can see that the relationships among the dimensions (Table 3) are not the same as students’ performance in the dimensions (Table 2). We make no attempts here to interpret differences in performance across disciplines; such differences may reflect variation in students’ preparation, variation in instructors’ teaching, disciplinary differences, etc. However, we are attempting to interpret the differences in relationships among the dimensions. While these differences, like the differences in performance, may have underlying causes related to students’ perspectives and instructors’ points of emphasis, we argue that those perspectives and points of emphasis reflect the very disciplinary epistemic beliefs that we are interested in better understanding. We recognize, of course, that this analysis is limited to three specific departments at Duke University.⁶ The arguments presented here resonate in conversations with instructors in these disciplines, but the perspectives and expectations of those instructors and the departments at this institution are not necessarily the same as those of the broader communities within each discipline. In short, more research is required to further test the ideas presented here.

While it may not seem groundbreaking to suggest that the social science of economics is more different from biology and chemistry than the two natural sciences are from one another, the fact that the same thesis assessment protocol is appropriate for three different disciplines implies that there is a shared orientation regarding the form of scholarly inquiry and communication. The protocol yields meaningful disciplinary rhetorical differences as well. Thus, we argue that the evidence increasingly suggests that one should not ask whether writing is generalized or discipline-specific but rather ask how to differentiate and teach that which is general and that which is discipline-specific.

This work carries implications and raises new questions that connect to both potential educational practice and future research. The implications for educational practices, particularly in WID programs, relate to the pedagogical strategy of making the links between genre conventions and discipline-specific values and ways of knowing visible to students. All of the dimensions that we discuss are valued across disciplines, but students may benefit from understanding how and why differences emerge. For example, it may be beneficial to make explicit to students why writing in biology involves a different kind of reasoning in the introduction than writing in chemistry, and how writing in chemistry builds that same kind of reasoning into the discussion. Instruction, whether occurring in writing studios or in writing courses to facilitate thesis preparation, may be more effective if it attends to disciplinary-specific as well as general elements of genre conventions, and the findings presented here lend empirical weight to conventional wisdom about how to discuss such genre conventions. We can only speculate about how effective these educational practices would be, but evidence suggests that they are worthy of investigation. Additionally, as our findings suggest that students' undergraduate honors theses reflect differences in disciplinary values and ways of knowing, the next step could be to conduct an analogous investigation of professional writing in these disciplines. Understanding if and how these patterns change as students become professionals (or, how professionals differ in their teaching of disciplinary writing and their practice of it) may shed further light on the development of scientists through writing.

The current study provides support for a quantitative method of genre analyses, TAP, that complements Wolfe et al.'s (2015) qualitative delineation of genre dimensions, particularly in terms of conceptual lens and macrostructures to construct arguments. This work provides evidence that within the genre of the honors thesis, *scientific reasoning in writing* is more strongly associated with *formulating a compelling argument for the significance of the research in the context of current literature* in biology, *making meaning regarding the implications of the findings* in chemistry, and *providing an organizational framework for interpreting the thesis* in economics.

Notes

1. Among biology theses, as reported by Reynolds and colleagues (2009), the Pearson correlation coefficient between raters' independent total scores, in which ratings on individual dimensions were combined, was 0.72. When scores were binned as either mastery (5) or non-mastery, the joint probability of agreement for individual dimensions ranged from 76% to 90%, with kappa values from 0.41 to 0.67 (all $p < 0.01$). Among economics theses, as reported by Dowd, Connolly, and colleagues (2015), the Pearson correlation coefficient between raters' independent scores was 0.94 for total thesis scores and ranged from 0.64 to 0.96 for the nine distinct dimensions. Among chemistry theses, as reported

by Dowd, Roy, and colleagues (2015), the Pearson correlation coefficient between raters' independent scores was 0.81 for total thesis scores and ranged from 0.53 to 0.78 for the nine distinct dimensions.

2. The test of factorial invariance (χ^2) compares two models in terms of criteria for best fit: the invariant model, in which factor loadings must be equal across all disciplines, and the variant model, in which factor loadings can vary freely in each discipline. The invariant model constrains the system because values that could be optimized independently of one another in the variant model now must be equal to one another. If the added constraint worsens the overall fit of the model, the χ^2 statistic will be statistically significantly greater and we can say that the factor loadings in the two disciplines differ. If the χ^2 statistic is not statistically significantly greater, we can say that the factor loadings in the two disciplines do not differ.

3. The shorthand names for TAP dimensions referenced in Tables 2 and 3 (and also referenced in the text) differ from shorthand names used in previous publications. We changed the names to better emphasize the nature of each dimension, but we emphasize that nothing about the actual dimensions or how they were assessed has changed.

4. In prior work, the differences between sub-groups within each discipline (students who participated in a specific intervention and students who did not) were investigated; those sub-groups are pooled together in both Table 2 and all of the analyses presented in this work.

5. Statistical analyses motivating the interpretation of some differences as being meaningful include both exploratory factor analyses of multiple bootstrapped samples from the data sets (which allows for comparison of confidence intervals for factor loadings) and confirmatory factor analyses in which only specific dimensions (as opposed to all dimensions) are constrained to be invariant. These analyses are not presented in this work, but they are available upon request.

6. We compared the underlying factor from biology theses at Duke University to that from biology theses collected at University of Minnesota between 2013 and 2015 to determine if the former was representative of biology as a discipline. Using the test of factorial invariance discussed in this work, we found that the factors are not statistically significantly different from one another. This supports our argument, although the support is certainly limited.

References

Carter, M. (2007). Ways of knowing, doing, and writing in the disciplines. *College Composition and Communication*, 58(3), 385–418.

- Dowd, J. E., Connolly, M. P., Thompson, Jr., R. J., & Reynolds, J. A. (2015). Improved reasoning in undergraduate writing through structured workshops. *The Journal of Economic Education*, 46(1), 14–27. doi: 10.1080/00220485.2014.978924
- Dowd, J. E., Roy, C. P., Thompson, Jr., R. J., & Reynolds, J. A. (2015). “On course” for supporting expanded participation and improving scientific reasoning in undergraduate thesis writing. *Journal of Chemical Education*, 92(1), 39–45. doi: 10.1021/ed500298r
- Herrington, A. J. (1981). Writing to learn: Writing across the disciplines. *College English*, 43(4), 379–87. doi: 10.2307/377126
- Kalman, C. (2009). The need to emphasize epistemology in teaching and research. *Science & Education*, 18(3–4), 325–47. doi: 10.1007/s11191-007-9135-1
- Kovac, J. (2003). Writing as thinking. *Annals of the New York Academy of Sciences*, 988(1), 233–38. doi: 10.1111/j.1749-6632.2003.tb06103.x
- Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science*, 67(1), 1–25.
- Morgan, D. L. (2007). Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research*, 1(1), 48–76. doi: 10.1177/2345678906292462
- Nussbaum, E. M., Sinatra, G. M., & Poliquin, A. (2008). Role of epistemic beliefs and scientific argumentation in science learning. *International Journal of Science Education*, 30(15), 1977–99. doi:10.1080/09500690701545919
- Phillips, D. C., & Burbules, N. C. (2000). *Postpositivism and educational research*. Lanham, MD: Rowman & Littlefield.
- Reynolds, J. A., Smith, R., Moskovitz, C., & Sayle, A. (2009). BioTAP: A systematic approach to teaching scientific writing and evaluating undergraduate theses. *BioScience*, 59(10), 896–903. doi: 10.1525/bio.2009.59.10.11
- Reynolds, J. A., & Thompson, R. J. (2011). Want to improve undergraduate thesis writing?: Engage students and their faculty readers in scientific peer review. *CBE-Life Sciences Education*, 10(2), 209–15. doi:10.1187/cbe.10-10-0127
- Robinson, M., Stoller, F., Costanza-Robinson, M., & Jones, J. K. (2008). *Write like a chemist: A guide and resource* (1st edition). Oxford, UK ; New York, NY: Oxford University Press.
- Rudolph, J. L. (2000). Reconsidering the “nature of science” as a curriculum component. *Journal of Curriculum Studies*, 32(3), 403–419. doi:10.1080/002202700182628
- Russ, R. S. (2014). Epistemology of science vs. epistemology for science. *Science Education*, 98(3), 388–96. doi: 10.1002/sce.21106
- Russell, D. R. (1997). Writing to learn to do: WAC, WAW, WAW—Wow! *Language and Learning Across the Disciplines*, 2(2), 3–8.
- Thagard, P. (2006). Conceptual change. In *Encyclopedia of cognitive science*. John Wiley & Sons, Ltd. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/0470018860.s00500/abstract>

- Van Dijk, E. M. (2014). Understanding the heterogeneous nature of science: A comprehensive notion of PCK for scientific literacy. *Science Education*, 98(3), 397–411. doi: 10.1002/sce.21110
- Wolfe, J., Olson, B., & Wilder, L. (2015). Knowing what we know about writing in the disciplines: A new approach to teaching for transfer in FYC. *The WAC Journal*, 25, 42–77.