JOURNAL OF CHEMICAL EDUCATION

Article

pubs.acs.org/jchemeduc

"On Course" for Supporting Expanded Participation and Improving Scientific Reasoning in Undergraduate Thesis Writing

Jason E. Dowd,*[†] Christopher P. Roy,[‡] Robert J. Thompson, Jr.,[§] and Julie A. Reynolds[†]

[†]Department of Biology, Duke University, P.O. Box 90338, Durham, North Carolina 27708, United States

[‡]Department of Chemistry, Duke University, P.O. Box 90354, Durham, North Carolina 27708, United States

[§]Department of Psychology and Neuroscience, Duke University, P.O. Box 90086, Durham, North Carolina 27708, United States

ABSTRACT: The Department of Chemistry at Duke University has endeavored to expand participation in undergraduate honors thesis research while maintaining the quality of the learning experience. Accomplishing this goal has been constrained by limited departmental resources (including faculty time) and increased diversity in students' preparation to engage in the research and writing processes. Here we assessed the relationship between iterative changes in pedagogical and mentoring support of honors research that efficiently employed departmental resources (including the chemistry thesis assessment protocol, ChemTAP) and students' scientific reasoning and writing skills reflected in their undergraduate theses. We found that, although we cannot disentangle some gradual changes over time from specific interventions, students exhibited the strongest performance when they participated in a course with structured scaffolding and used assessment tools explicitly designed to enhance the scientific reasoning in writing. Furthermore, less prepared students exhibited more positive changes.

KEYWORDS: *General Public, Communication/Writing, Undergraduate Research, Chemical Education Research*

■ INTRODUCTION

In 1998, the Boyer Commission on Educating Undergraduates in the Research University identified a number of deficiencies in undergraduate education at research universities, including limited student engagement with sophisticated, collaborative major research endeavors.[1] In response, a common goal for many postsecondary institutions has been to expand participation in undergraduate research. The research laboratory more closely models the professional setting within the sciences than the classroom, and therefore offers students experiences that they would not otherwise have as undergraduates.[2] Moreover, the experience of working closely with mentors affords the opportunity for both explicit and tacit learning experiences through apprenticeship,[3,4] bridging the gap between expert and novice ways of approaching problem solving and critical thinking within the discipline.[5,6] Undergraduate research is recognized as a high-impact practice,[7] and evidence indicates that participating in the research process improves undergraduate science students' skills with regard to critical thinking, problem solving, and applying knowledge and engenders a more sophisticated understanding of the nature and construction of knowledge.[8,9]

In particular, capstone research experiences emphasize students' independence as they carry out complete projects and submit undergraduate honors theses. Although the research and writing processes may appear to be distinct, they cannot be disentangled. Writing may occur after data collection and analysis are complete, but students must continue to consider and synthesize aspects of their research throughout the writing process, which serves to make thinking visible.

The goal of expanding participation in such capstone research experiences is constrained, however, in two ways: (1) limited departmental resources (including faculty time); and (2) increased diversity in students' preparation to engage in the research and writing processes. Preparation of this expanded cohort of students refers not only to traditional markers such as math skills and previous independent study experience, but also to epistemological beliefs, intrinsic motivation, and metacognitive skills. Furthermore, variations in how regularly faculty advisors meets with students, how much constructive feedback is provided, and numerous other factors may lead to learning experiences that vary substantially in quality. As colleges and universities continue to encourage more students with more widely varying backgrounds to participate in undergraduate research experiences, there is a pressing need to ensure that the quality of the experiences does not decline. Therefore, efforts to expand participation require not only efficient use of departmental resources, but also effective pedagogical approaches and mentorship.

Previous work has demonstrated that structured thesis-writing courses intended to promote metacognition can improve both writing skills and critical thinking skills, allowing departments to increase participation rates while simultaneously increasing the quality of student work.[10,11] In these courses, students learn the conventions of professional scientific peer review and how to interact with their faculty research mentors in a more professional manner, and the efficiency of those interactions is increased through the use of objective,

dx.doi.org/10.1021/ed500298r | J. Chem. Educ. XXXX, XXX, XXX−XXX

outcome-oriented assessment tools.[11] In our university's biology department, we saw an increase in the number of students who wrote a thesis during the years in which these changes were implemented, and the quality of scientific reasoning in students' writing was significantly stronger when students used the scaffolding and writing assessment tools in the context of a thesis-writing course.

Pedagogical practices that work well in one course, department, or institution may not necessarily work as well in another. A major challenge of educational research is developing a conceptual framework that helps us understand why certain practices are effective within specific educational contexts.[12] Nonetheless, discipline-based science education researchers have shown that some practices can be effective across contexts (e.g., active learning in large introductory classes).[13−16] Here, we evaluate the adaptation and incorporation of pedagogical approaches developed in one context (the biology department) into a new context, an existing honors chemistry course within the same university. In doing so, we intend to contribute to the broader discussion of students' writing as a window to scientific thinking and reasoning. This course supports students completing an undergraduate honors thesis by scaffolding and facilitating the thesis-writing process while efficiently expanding faculty resources for undergraduate research. In this analysis, we pose the question: What is the relationship between changes in the support of undergraduate thesis writing and students' scientific reasoning and writing skills? To address this question, we analyzed a sample of student theses completed between 2001 and 2013 in the Department of Chemistry at Duke University.

Context

Prior to 2007, students completing an undergraduate thesis in chemistry had no explicit departmental guidelines or requirements other than a minimum GPA requirement. In formulating their capstone research experience, students consulted with faculty members and joined research groups on an individual basis. Theses were evaluated for graduation with distinction by a committee of faculty members after submission; there was no formative feedback or standardized means of assessment during the writing process.

The University implemented a new general education curriculum in 2000 that required students to complete three writing-intensive courses and two research-oriented courses, and also provided departments with support for summer research opportunities for undergraduates. In response to these initiatives to encourage more undergraduates to engage in research, the number of students writing honors theses increased from year to year and tightened constraints on faculty time and resources. Some of the practices and disciplinary understandings in chemistry are common across research areas, and therefore are more efficiently addressed with students in a group setting. Beginning in 2007, students who met the GPA requirement and wanted to graduate with distinction in chemistry were not only required to complete an undergraduate thesis but also to participate in a sequence of two half-credit honors capstone courses. The first course, taken concurrently with the first semester of thesis-related independent study, focused on safety procedures, properly maintaining a laboratory notebook, and ethical behavior in research. It culminated in the writing of a research proposal, setting the stage for subsequent research. In the second course, students learned the conventions of scientific writing, practiced constructing a concise narrative from their research, and attuned themselves to readers' expectations through soliciting and responding to feedback.

During the second of the two courses, early meetings were devoted to introducing the course goals and explaining library and bibliographic resources. Instructors then provided students with an overview of each part of the thesis and students peer-reviewed drafts of each other's work. Students also presented research in class and engaged in peer review of these presentations. As the course instructor was also a primary evaluator of whether students' theses ultimately merited graduation with distinction, the instructor did not provide specific suggestions to individual students throughout the semester. Instead, students regularly solicited feedback from their research mentors and also provided feedback to one another. Toward the end of the semester, class meetings focused on specific aspects of students' writing (e.g., conciseness in writing). Students presented final oral presentations and posters to their peers. Course grades were based on the final draft of the thesis, as well as peer reviews and in-class participation. A committee composed of the student's research supervisor, the course instructor, and one additional faculty reader determined whether students graduated with distinction.

In 2010, the chemistry thesis assessment protocol (ChemTAP) was introduced into the second honors capstone course. ChemTAP, which was adapted from the closely related biology thesis assessment protocol,[11] was designed to both focus students on the most important aspects of writing a thesis and allow for objective, transparent assessment of the work. With the use of ChemTAP, students reviewed each other's writing and responded to one another just as professional academics would in the peer review process, though in a more scaffolded manner. ChemTAP does not solicit "peer editing" or "peer grading," and students do not comment on the accuracy or appropriateness of the research. Instead, students use the rubric to address higher-order critical thinking skills, such as those reflected in constructing a scientific argument. Although faculty mentors certainly have more research-related expertise than students' peers, peers are an appropriate audience for the assessment of scientific reasoning in research. If these practices are modeled in class by the instructor, students can do remarkably well. The peer review worksheets are very similar to those available in ref 11. The protocol involved implementation of rubrics in classroom activities, peer review, and faculty evaluation.

Both the sequence of courses and the implementation of ChemTAP are implicitly rooted in the social constructivist theory of knowledge and learning.[17] Particularly in ChemTAP, activities and resources are designed for students to consider their pre-existing ideas, confront inconsistencies, and ultimately resolve those inconsistencies for improved scientific reasoning in writing. This process occurs through social interactions, which are facilitated by the classroom setting and resources.

■ METHODS

Study Sample

From 2000 to 2013, 239 students submitted an undergraduate honors thesis and graduated with distinction in chemistry. The number of undergraduates who completed honors theses generally increased over the years ($r = 0.54$, $p = 0.044$; Figure 1). There was a decline in chemistry majors in 2013 (which we

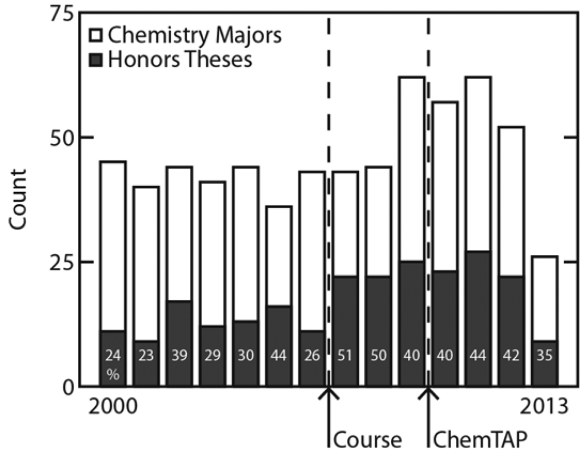

Figure 1. Number of graduates and the fraction of graduates completing honors theses (displayed within each column) generally increase over this time period.

speculate results from the introduction of neuroscience as a major at Duke University), but the proportion of chemistry majors choosing to complete an honors thesis also increased over time ($r = 0.56$, $p = 0.038$; Figure 1).

With the increase in number of students participating in undergraduate research over the years, the educational experiences of those students also changed and increasingly varied. To numerically characterize such changes, background characteristics were averaged by year, resulting in up to 14 mean and standard deviation values (2000−2013). The average numbers of both research-oriented and writing-intensive courses completed by students increased over time ($r = 0.80$, $p = 0.0006$, and $r = 0.76$, $p = 0.0017$, respectively; Table 1). The

Table 1. Pairwise Correlations between Student Population Characteristics and Both Time and Number of Honors Graduates

Characteristic	Year[a]	Honors Graduates[a]
Independent Study Courses (μ)	0.50	0.70**
Independent Study Courses (σ)	0.37	0.60*
Research-Oriented Courses (μ)	0.80***	0.63*
Research-Oriented Courses (σ)	0.36	0.51
Writing-Intensive Courses (μ)	0.76**	0.46
Writing-Intensive Courses (σ)	0.14	0.48

[a]*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

standard deviation in the number of independent study courses completed by students increased with the increasing number of students completing honors theses ($r = 0.60$, $p = 0.025$; Table 1). These factors reflect the changing educational context and student population characteristics to which the chemistry department was adapting and responding.

These changes over time were paralleled in the Department of Chemistry by the additional scaffolding, honors thesis capstone course and the incorporation of ChemTAP, all of which reflect a trajectory explicitly designed to enhance scientific reasoning in writing. The average numbers of independent study, research-oriented, and writing-intensive courses completed by students were significantly greater during years in which students participated in the thesis writing course than during prior years ($p < 0.01$ in all comparisons), according to Student's *t*-tests (Table 2). However, there are no statistically significant differences in the number of independent study, research-oriented, and writing-intensive courses when comparing students participating in the thesis writing course before and after the implementation of ChemTAP in 2010.

To control for individual differences in skills and experiences prior to writing a thesis, we obtained the following background information (when available) for students through the university registrar, as approved by our Institutional Review Board: GPA at graduation ($n = 237$), SAT scores ($n = 231$), grade in the first-year writing course ($n = 236$), number of independent study courses taken ($n = 237$), number of research-based courses taken ($n = 237$), number of writing-intensive courses taken ($n = 237$), total number of courses taken ($n = 237$), and gender ($n = 235$). For each of our analyses, we included as many students as possible (for example, we did not exclude a student for whom we did not have gender information from an analysis of how students' research-oriented course experience changes over the years).

To address the question of the relationship between changes in the support of undergraduate thesis writing and students' scientific reasoning and writing skills, we decided to evaluate a sample of the 239 theses. The only selection factor was to ensure that the subset included theses reflecting each of the three intervention phases (no course, course, and course with ChemTap). The resulting subset included 93 theses. To determine if our sample reflects the total population, we compared the background characteristics of assessed versus unassessed students to see if they were significantly different. We found that students whose theses were assessed had taken fewer research-based courses than students whose theses were not assessed ($\mu_{scored} = 3.60$, $\mu_{unscored} = 4.17$, $p = 0.05$). Our sample under-represents males during the time frame of 2007−2009 ($male_{scored} = 28.6\%$, $male_{unscored} = 59.3\%$, $p = 0.04$), and students in our sample completed more writing-intensive courses than students whose theses were not assessed during the time frame of 2010−2013 ($\mu_{scored} = 2.65$, $\mu_{unscored} = 2.19$, $p = 0.03$). Even though differences in mean values do not necessarily correspond to differences in relationships among

Table 2. Descriptive Statistics of Student Population Characteristics, Separated by Each Intervention

Characteristic	No Course ($N = 89$)	Course ($N = 68$)	ChemTAP ($N = 80$)	*p*-value[a]
Independent Study Courses (μ)	2.60	3.08	3.11	0.0012
Independent Study Courses (σ)	1.06	1.15	1.19	-
Research-Oriented Courses (μ)	2.67	4.65	4.77	<0.0001
Research-Oriented Courses (σ)	2.14	1.85	1.69	-
Writing-Intensive Courses (μ)	1.59	2.48	2.33	<0.0001
Writing-Intensive Courses (σ)	1.27	1.17	0.88	-

[a]The *p*-values shown here refer to differences between students who did not participate in the thesis writing course and students who participated in the course (with or without ChemTAP).

factors, such differences indicate that missing values may not be randomly distributed. We must be cautious about generalizing analyses of the assessed sample involving these factors to the larger population.

The quality of the research, the laboratory setting, and the interaction with the advisor play an important role in thesis development. Such aspects may influence students' scientific reasoning exhibited in their writing. While we cannot control for variation in these factors, we have no reason to believe that they are changing over time on average. Additionally, students are encouraged to write honors theses even if the results of the research are inconclusive or limited, as long as they present, analyze, and discuss the implications of them as such; so an unsuccessful research experience may lead to an effective, well-reasoned thesis.

Chemistry Thesis Assessment Protocol (ChemTAP)

ChemTAP, just like the biology thesis assessment protocol, is a document that guides and supports students and instructors through the thesis-writing process.[11] The ChemTAP rubric contains 13 distinct dimensions assessing students' writing skills, scientific reasoning, and accuracy and appropriateness of research. Detailed descriptions of these dimensions, which apply to both chemistry and biology, are available in ref 11. In this study, we focused exclusively on the nine dimensions related to writing and reasoning, as the expertise required to evaluate the accuracy and appropriateness of research across so many subdisciplines within chemistry was not available within the scope of this study. Specifically, we assessed theses in terms of: appropriateness for target audience, argument for significance of research, articulation of goals, interpretation of results, implications of findings, organization, absence of writing errors, consistent and professional citations, and effective use of tables and figures.

Each dimension was rated on a scale of 1 to 5. A rating of 1 indicates that the dimension under consideration is either missing, incomplete, or below the minimum acceptable standards. A rating of 3 indicates that the dimension is adequate, but the work does not exhibit mastery. A rating of 5 indicates that the dimension is excellent and the work exhibits mastery. As different parts of the thesis might fall into different categories, intermediate ratings of 2 and 4 may be appropriate.

Theses written between 2000 and 2012 were assessed by two of the authors (J.A.R. and C.P.R.) over the course of one semester. Over 1 week in 2013, support from an NSF grant allowed us to hire and train a group of graduate students and postdoctoral fellows, including one author (J.E.D.), to assess the 2013 theses. For all assessment, each rater completed training in the use of the ChemTAP rubric, which included examination of samples of students' writings that illustrated inadequate, adequate and masterful levels of all nine dimensions being assessed. Raters then assessed sample theses that were not part of the data set, discussed them, and established consensus scores as a means of calibrating. Raters of 2013 theses were trained and calibrated using a sample of theses assessed from 2000 to 2012.

Each thesis in our sample was read by two raters who assessed the theses independently, subsequently discussed discrepancies in their ratings, and, finally, established a consensus score (the same method employed in ref 10). The consensus score was not merely an average, but a discussion-based final score agreed upon by both raters. The Pearson correlation coefficient between raters' independent scores is 0.81 for total thesis scores and range from 0.53 to 0.78 for the nine distinct dimensions. Consensus scores were used in all analyses.

Formally, the nine dimensions assessed using the ChemTAP rubric are independent of one another. Because there is no *a priori* reason to consider mastery in argument, for example, on the same scale as mastery in employing figures and tables, the total score is not necessarily meaningful. However, exploratory factor analysis indicated that there is a latent underlying factor representing students' exhibition of scientific reasoning in chemistry thesis writing. The factor loadings revealed that the ability to interpret results and draw implications loads most strongly onto the factor, with organization, establishment of argument and identification of audience also loading strongly (Table 3). We considered this factor the primary characteristic of students' exhibition of scientific reasoning in chemistry thesis writing.

Table 3. Factor Loadings of Nine Dimensions of Students' Theses on the Dominant Underlying Factor

Dimension	Factor Loadings
Identifying audience	0.4574
Structuring argument	0.4951
Stating goals	0.4015
Interpretation of results	0.6484
Implications of results	0.8050
Organization	0.5472
Minimizing writing errors	0.3340
Appropriate citation	0.3810
Effective tables and figures	0.2526

We investigated the relationship between this thesis assessment factor and two interventions: the introduction of the thesis writing course in 2007 and the introduction of ChemTAP in 2010. Using linear regression analysis, we were able to control for numerous covariates that could relate to thesis assessment. These covariates are GPA at graduation, SAT scores, grade in the first-year writing course, number of independent study courses taken, number of research-based courses taken, number of writing-intensive courses taken, total number of courses taken, and gender. Finally, we used linear regression analysis to explore potential relationships between these covariates and the thesis assessment factor both within and across each intervention; such relationships could indicate whether specific subsets of students exhibit stronger scientific reasoning in thesis writing than others.

■ RESULTS

Students who participated in the course when ChemTAP was implemented exhibited stronger scientific reasoning in their thesis writing than students in the course without ChemTAP ($\beta = 0.365$, $p = 0.014$), according to linear regression analysis (Table 4, Figure 2). None of the additional covariates were statistically significant.

However, when the interaction between additional covariates and interventions was considered in regression models, the most parsimonious linear regression model for predicting the primary thesis assessment factor included the introduction of the writing course, the introduction of ChemTAP, students' SAT math scores, the number of research-oriented courses, and the interaction between SAT scores and research-oriented

Table 4. Fitted Linear Regression Model Predicting Thesis Assessment by Thesis Writing Experience[a]

Parameter	Model 1[b,c]
Course	−0.179
ChemTAP	0.365*

[a]N = 93. All coefficients are shown here in standard units. [b]Adjusted R^2 = 0.0514 in Model 1. [c]*p < 0.05, **p < 0.01, ***p < 0.001.

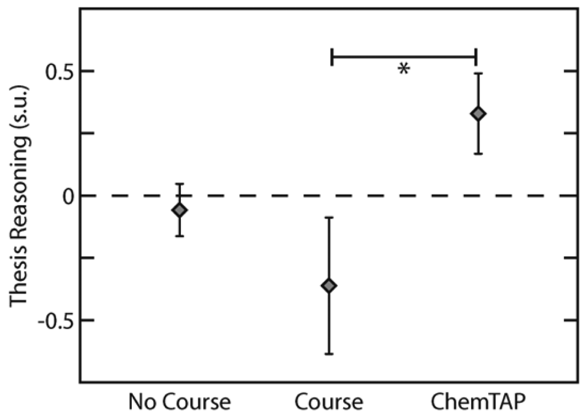

Figure 2. Assessment of students' scientific reasoning expressed through writing in the undergraduate honors thesis ("Thesis Reasoning," shown in standard units) is shown. The asterisk indicates that the comparison is statistically significant at the p = 0.05-level.

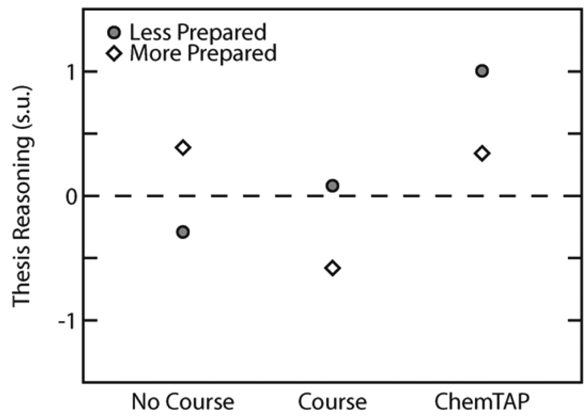

Figure 3. Predicted assessment of students' scientific reasoning expressed through thesis writing ("Thesis Reasoning," shown in standard units) of students with lowest quartile SAT math scores and numbers of research courses (Less Prepared) and students with highest quartile SAT math scores and numbers of research courses (More Prepared) are shown.

courses and the introduction of the thesis writing class (Table 5).

In Model 1, the implementation of ChemTAP is associated with a statistically significant positive change in thesis assessment. In Model 2, the main effects of SAT math scores and research-oriented course experience are included, though they are not statistically significantly related. In Model 3, almost all of the variables are statistically significant with the inclusion of the two interaction terms (Figure 3). In this model, the implementation of ChemTAP is associated with a positive change in thesis assessment. Both interaction terms are negative and the magnitude of each term is larger than its respective main effect, indicating that participation in the thesis writing course is associated with more positive changes in thesis assessment for students with lower SAT math scores and fewer research courses (whom we describe as "less prepared") than for students with higher SAT math scores and more research courses (whom we describe as "more prepared"). In other words, the difference in thesis assessment between more and less prepared students that is present before any intervention disappears when the thesis writing course is introduced.

Regression models in Table 5 are based on a subsample of 88 students for whom we had SAT scores, course information, and thesis assessment data. During the time frame of 2010–2013, the students included in this regression model exhibited stronger thesis performance than students for whom we did not have SAT scores and course information ($\mu_{complete}$ = 0.49, $\mu_{incomplete}$ = −0.84, p = 0.005). This difference amplified the positive change in performance associated with the introduction of ChemTAP, which is why the relationship in Model 1 of Table 5 is stronger than the relationship in Table 4 exhibited by the sample of 93 students whose theses were assessed.

During the years in which students wrote honors theses without a writing course (2000–2006), we observed a positive correlation between incoming SAT math score and overall thesis assessment (r = 0.40, p = 0.003), as well as a positive correlation between the total number of research-oriented courses taken and overall thesis assessment (r = 0.32, p = 0.017). Linear regression analysis indicated that both of these relationships explain variation in students' thesis assessment (Table 6), which is consistent with the main effects in Model 3 of Table 5.

During subsequent years in which students wrote honors theses while participating in a writing course but without the additional structure of ChemTAP (2007–2009), we observed no statistically significant relationships between overall thesis assessment and gender, grade point average, incoming SAT scores, class rank, number of independent study courses completed, number of research courses taken, number of

Table 5. Fitted Linear Regression Models Predicting Thesis Assessment by Selected Variables Representing Students' Background and Thesis Writing Experience[a]

Parameter	Model 1[b,e]	Model 2[c,e]	Model 3[d,e]
Course	−0.197	−0.258	3.792**
ChemTAP	0.459**	0.459**	0.474**
SAT Math		0.137	0.298*
Research-Oriented Courses		0.096	0.257
SAT Math * Course			−3.513*
Research-Oriented Courses * Course			−0.709*

[a]N = 88. All coefficients are shown here in standard units. [b]Adjusted R^2 = 0.100 in Model 1. [c]Adjusted R^2 = 0.110 in Model 2. [d]Adjusted R^2 = 0.224 in Model 3. [e]*p < 0.05, **p < 0.01, ***p < 0.001.

Table 6. Fitted Linear Regression Models Predicting Thesis Assessment by Selected Variables Representing Student Background Characteristics from 2000 to 2006[a]

Parameter	Model 1[b,e]	Model 2[c,e]	Model 3[d,e]
SAT Math	0.396**		0.348**
Research-Oriented Courses		0.331*	0.268*

[a] $N = 53$. All coefficients are shown here in standard units. [b] adjusted $R^2 = 0.141$ in Model 1. [c] adjusted $R^2 = 0.092$ in Model 2. [d] adjusted $R^2 = 0.196$ in Model 3. [e] *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

writing courses taken, or freshman writing course grade. We also observed no statistically significant relationships between overall thesis assessment and these other factors when ChemTAP was implemented (2010−2013).

DISCUSSION

The Department of Chemistry adapted to both the increasing number and diversity of experiences of students participating in research through iterative changes in pedagogical structure and mentoring that were rooted in the social constructivist theory of knowledge and learning.[17] In this theoretical framework, the scaffolding for students' interactions with peers and opportunities to reflect on their own understanding provide the tools for improved scientific reasoning in writing. Evaluation of the effectiveness of these adaptations with regard to scientific reasoning skills is critical for continued efficient use of resources for improving educational practices, as improved scientific reasoning is an important outcome of participating in undergraduate research. We observed a statistically significant increase in students' scientific reasoning in their thesis writing when ChemTAP was implemented in 2010 (Table 4 and Figure 2). Research involving evaluation of effectiveness does not employ experimental control of the multiple potentially contributing factors and, therefore, does not seek to establish a causal mechanism. Although we cannot claim that the implementation of ChemTAP is causally responsible for increased reasoning, the evidence regarding effectiveness is encouraging.

Prior to the implementation of the writing course in 2007, students who completed more research-oriented courses in college and scored better on the math portion of the SAT (which we treated as a proxy for incoming mathematical ability) exhibited stronger scientific reasoning in thesis writing (Table 6). However, with the implementation of the writing course (both without and with ChemTAP), scientific reasoning in thesis writing was no longer related to these background characteristics. Furthermore, once the thesis writing course was implemented, less prepared students exhibited more positive changes in scientific reasoning than the more prepared students (Table 5 and Figure 3). These findings suggest not only that students' scientific reasoning in chemistry thesis writing improved overall with the increased structure provided by the course with ChemTAP, but also that the implementation of the course had a larger impact on the performance of the less prepared students. This observation is particularly important because variation in some aspects of students' preparation increases with increasing participation (Table 1).

On average, students who participated in the thesis writing course completed more independent study, research-oriented and writing-intensive classes than students who did not participate in the course (Table 2). However, the initial implementation of the writing course in 2007, while accommodating the increase in number and diversity of experiences of students, is not associated with improved scientific reasoning. Instead, we observed improved scientific reasoning only upon the implementation of ChemTAP. We observed no differences in students' average background when we compared students who participated in the thesis writing course before and after the implementation of ChemTAP in 2010. Regression models show that none of the background variables were associated with stronger thesis assessment after the writing course was implemented. Thus, although students had stronger incoming experiences and were, perhaps, better equipped to undertake thesis research during the years in which the thesis writing course was implemented, these differences do not explain the improved performance. Because ChemTAP explicitly encourages students to think about and discuss scientific reasoning in their writing, we hypothesize that its introduction into the seminar course provided students with critical, highly structured scaffolding to promote metacognition and improve learning outcomes, in keeping with the mechanisms for learning characterized in the social constructivist theory of knowledge and learning.[17] We are testing this hypothesis in ongoing research that involves multiple departments at Duke University and several other universities.

Although the study presented here is tightly focused on one department at one institution, it also contributes to a larger effort to study students' writing as a window to scientific thinking and reasoning.[12] Our results not only parallel those found in the biology department's thesis writing course,[10,11] but they are also similar to findings from other institutions and other contexts. For example, in both introductory and advanced chemistry courses, increased structure for writing laboratory reports was associated with improved performance.[18,19] Similarly, in introductory biology courses, increased structure in active learning practices is associated with improved performance in reasoning tasks, particularly among students with disadvantaged educational backgrounds.[14] Even though the student populations across all these studies differ, we highlight the consistent link between (1) increased scaffolding for metacognitive engagement, (2) improved scientific reasoning, and (3) disproportionate gains by students most at risk.

Limitations

The limitations of this study are related to the lack of experimental control inherent in research that aims to evaluate the effectiveness of educational practices involving multiple components that operate concurrently. The implementation of ChemTAP is not the only explanation for positive changes in students' scientific reasoning exhibited in their theses. The same instructors led the class throughout all the years, which may suggest that increased experience and familiarity with the course improved the measured scientific reasoning. Adaptations in materials and activities, such as improved use of peer review among students, surely contributed to year-to-year changes. Unfortunately, we could not control for subsequent years in a linear regression model alongside the implementation of the thesis writing course and ChemTAP because all of these factors represent changes over time and are highly correlated with one another.

Additionally, the student population changed over the years. Although we found that few of the factors for which we have information were statistically significant, our ability to control for incoming ability in scientific reasoning and writing was limited by the scholastic factors to which we have access. It is possible that unmeasured differences (quality of the research experience, the number of semesters of *thesis-related* research, the number of *technical* writing courses, GPA in STEM courses, or postgraduation plans, to name a few) played a role, and the effect of these specific interventions was diminished by other changes.

Nonetheless, these results suggest that the incorporation of a thesis writing course was among changes positively associated with relative gains of less-prepared students, and the incorporation of ChemTAP was among changes positively associated with improved average exhibition of scientific reasoning in the thesis. Despite limitations, interventions in the capstone research and thesis writing experience in chemistry appear to accommodate larger student populations without forfeiting strong learning outcomes.

■ CONCLUSION

After several years of increasing participation of students with increasingly varied degrees of preparation in undergraduate honors thesis research, we found that changes to the support of this research are, indeed, related to the assessment of students' scientific reasoning and writing skills reflected in their written undergraduate thesis. Although we could not disentangle some gradual changes over time from specific interventions, students who participated in structured scaffolding and employed assessment tools explicitly designed to enhance the scientific reasoning in writing seemed to exhibit the strongest performance. The chemistry department is successfully engaging students and improving the pedagogical resources available to them. Although differing contexts of departments, institutions and student populations present a challenge to making generalizations based on these findings, the connections that we have highlighted with other efforts in discipline-based education research suggest that the implementation of ChemTAP, an intervention specifically designed to engage students' metacognitive skills, is a positive step for efficiently increasing and enhancing participation in honors research. We hope that educators may apply these resources in their own settings, and that this work ultimately stimulates the chemical education research community to further investigate writing as a window to scientific reasoning.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: jason.dowd@duke.edu.

Author Contributions

Several people contributed to the work described in this paper. J.A.R., C.P.R., and J.E.D. collected data and assessed students' theses. J.E.D. conducted subsequent data analysis, in close collaboration with J.A.R. and R.J.T, Jr. J.A.R. supervised the research and the development of the manuscript. J.E.D. wrote the first draft of the manuscript; all authors subsequently took part in the revision process and approved the final copy of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The research described in this paper was supported in part by the National Science Foundation under award 1225768.

■ REFERENCES

(1) Kenny, R. W.; Boyer Commission on Educating Undergraduates in the Research University. *Reinventing Undergraduate Education a Blueprint for America's Research Universities*; Carnegie Foundation for the Advancement of Teaching: New York, 1998.

(2) Shulman, L. S. Signature Pedagogies in the Professions. *Daedalus* **2005**, *134* (3), 52−59.

(3) Hennessy, S. Situated Cognition and Cognitive Apprenticeship: Implications for Classroom Learning. *Stud. Sci. Educ.* **1993**, *22* (1), 1−41.

(4) Choi, J. I.; Hannafin, M. Situated Cognition and Learning Environments: Roles, Structures, and Implications for Design. *Educ. Technol. Res. Dev.* **1995**, *43* (2), 53−69.

(5) Chi, M. T. H.; Feltovich, P. J.; Glaser, R. Categorization and Representation of Physics Problems by Experts and Novices. *Cognit. Sci.* **1981**, *5* (2), 121−152.

(6) Hardiman, P. T.; Dufresne, R.; Mestre, J. P. The relation between problem categorization and problem solving among experts and novices. *Mem. Cogn.* **1989**, *17* (5), 627−638.

(7) Kuh, G. D.; Schneider, C. G.; Association of American Colleges and Universities. *High-Impact Educational Practices: What They Are, Who Has Access to Them, and Why They Matter*; Association of American Colleges and Universities: Washington, DC, 2008.

(8) Lopatto, D. Survey of Undergraduate Research Experiences (SURE): First Findings. *Cell Biol. Educ.* **2004**, *3* (4), 270−277.

(9) Seymour, E.; Hunter, A. B.; Laursen, S. L.; DeAntoni, T. Establishing the Benefits of Research Experiences for Undergraduates in the Sciences: First Findings from a Three-Year Study. *Sci Educ.* **2004**, *88* (4), 493−534.

(10) Reynolds, J. A.; Thompson, R. J. Want To Improve Undergraduate Thesis Writing? Engage Students and Their Faculty Readers in Scientific Peer Review. *CBE Life Sci. Educ.* **2011**, *10* (2), 209−215.

(11) Reynolds, J.; Smith, R.; Moskovitz, C.; Sayle, A. BioTAP: A Systematic Approach to Teaching Scientific Writing and Evaluating Undergraduate Theses. *BioScience* **2009**, *59* (10), 896−903.

(12) Reynolds, J. A.; Thaiss, C.; Katkin, W.; Thompson, R. J. Writing-to-Learn in Undergraduate Science Education: A Community-Based, Conceptually Driven Approach. *CBE Life Sci. Educ.* **2012**, *11* (1), 17−25.

(13) Freeman, S.; O'Connor, E.; Parks, J. W.; Cunningham, M.; Hurley, D.; Haak, D.; Dirks, C.; Wenderoth, M. P. Prescribed Active Learning Increases Performance in Introductory Biology. *CBE Life Sci. Educ.* **2007**, *6* (2), 132−139.

(14) Haak, D. C.; HilleRisLambers, J.; Pitre, E.; Freeman, S. Increased Structure and Active Learning Reduce the Achievement Gap in Introductory Biology. *Science* **2011**, *332* (6034), 1213−1216.

(15) Fraser, J. M.; Timan, A. L.; Miller, K.; Dowd, J. E.; Tucker, L.; Mazur, E. Teaching and Physics Education Research: Bridging the Gap. *Rep. Prog. Phys.* **2014**, *77* (3), 032401.

(16) Crouch, C. H.; Mazur, E. Peer Instruction: Ten Years of Experience and Results. *Am. J. Phys.* **2001**, *69* (9), 970−977.

(17) Duit, R.; Treagust, D. F. Learning in Science—From Behaviourism towards Social Constructivism and Beyond. In *International Handbook of Science Education, Part 1*; Fraser, B. J., Tobin, K., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1998; pp 3−25.

(18) Deiner, L. J.; Newsome, D.; Samaroo, D. Directed Self-Inquiry: A Scaffold for Teaching Laboratory Report Writing. *J. Chem. Educ.* **2012**, *89* (12), 1511−1514.

(19) Van Bramer, S. E.; Bastin, L. D. Using a Progressive Paper To Develop Students' Writing Skills. *J. Chem. Educ.* **2013**, *90* (6), 745−750.