JOURNAL OF CHEMICAL EDUCATION
Article
pubs.acs.org/jchemeduc

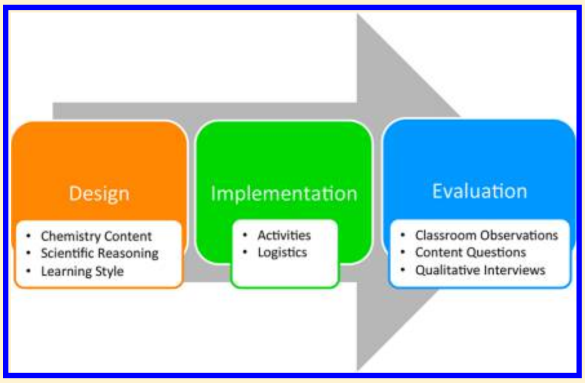
Targeting the Development of Content Knowledge and Scientific Reasoning: Reforming College-Level Chemistry for Nonscience Majors

Justin H. Carmel, Yasmin Jessa, and Ellen J. Yezierski*

Department of Chemistry and Biochemistry, Miami University, 701 East High Street, Oxford, Ohio 45056, United States

Supporting Information

ABSTRACT: A liberal education curriculum requires discipline-specific courses that develop intellectual and practical skills. With this promise of development, it is crucial that instruction focuses on content knowledge as well as the thinking patterns associated with the content. In chemistry, scientific reasoning is one such skill that students should improve; however, this is a tall order for a traditional lecture course. Through the use of student-centered, content-driven, and reasoning skill-focused interventions, this development may be more likely to occur. This article describes the development, implementation, and evaluation of instructional materials for a chemistry class for non-science majors and makes the materials available to the chemistry education community for broader dissemination.



KEYWORDS: *Curriculum, First-Year Undergraduate/General, Collaborative/Cooperative Learning, Hands-On Learning/Manipulatives, Problem Solving/Decision Making, Nonmajor Courses*

BACKGROUND

Many universities across the country subscribe to the idea of liberal education; their students take a variety of core courses to ensure exposure to content and theories outside of their major field to help "develop both intellectual and practical skills."[1] Of the practical skills, the one that is given the most emphasis is critical thinking. Critical thinking as a construct has a notoriously broad range of definitions, from "applying, analyzing, synthesizing, and/or evaluating information"[2] to "purposeful, self-regulatory judgment,"[3] making it difficult to assess and thus target instruction and measure improvement. We selected scientific reasoning (SR) as a construct of interest since it comprises a subset of critical thinking skills that fits both criteria, as it focuses on examining data, making inferences, and hypothesizing, as well as basic reasoning tasks. Previous literature has shown that improvement in SR can be accomplished through the use of inquiry-based learning approaches.[4,5] In our efforts to improve student content and SR outcomes, we set out to design and evaluate classroom interventions for a chemistry course for nonscience majors. This article describes (1) the development of content and SR-focused classroom interventions; (2) course specifics and implementation; and (3) evaluation methods and outcomes. While incorporating student-centered instruction in the classroom is not a novel concept, we have taken extra care to not only create interventions but to also evaluate their intended design features using a combination of methods commonly used by researchers, which is uncommon in dissemination of classroom materials. While collaborative pencil-and-paper activities have been created for courses such as general chemistry,[6,7] analytical chemistry,[8] and physical chemistry,[9,10] reformed classroom materials for this particular chemistry course are not typically found. We hope that the interventions and data to describe their effectiveness will motivate other faculty to use them, as well as provide useful methods for further curriculum development and testing. The interventions were developed as part of a larger study that investigates the effect of targeted instruction on the students' scientific reasoning skills. Data describing students' reasoning gains without the use of the interventions has been previously published.[6]

DESIGN

The interventions were designed for use in a chemistry class for nonscience majors at our home institution. The majority of the students enrolled in this course are business majors with smaller contingents from the humanities and education, which is typical for this course. The average self-reported ACT score for these students is 26, which is commensurate with the profile of the entire university. The course, Chemistry in Modern Society, aims to help students understand both the nature of basic chemical processes and the ways that chemistry affects our society. The course text is the American Chemical Society's Chemistry in Context (7th ed.),[12] and the interventions were

dx.doi.org/10.1021/ed500207t | J. Chem. Educ. XXXX, XXX, XXX−XXX

developed to be seamlessly integrated into the class and in parallel with the textbook. The course addresses material from five chapters of the textbook (Table 1). Also shown are the specific content foci of the interventions.

Table 1. Textbook Chapters Taught and Intervention Content Foci

Textbook Chapter	Content Focus
Ch. 1: The Air We Breathe	The Air Quality Index
Ch. 2: Protecting the Ozone Layer	Properties of Light
Ch. 3: The Chemistry of Global Warming	CO_2 and The Greenhouse Effect
Ch. 7: The Fires of Nuclear Fission	Radioactive Half-Lives and Chernobyl
Ch. 11: Nutrition: Food for Thought	Molecular Polarity[a] and Food

[a]Note: Polarity is presented in Ch. 5 but is taught in conjunction with Ch. 11.

The interventions are intended to take between 45 and 50 min to complete—a typical lecture period for a three credit hour course. Students are placed in heterogeneous groups of 4–5 students based on their score on the Lawson Classroom Test of Scientific Reasoning (LCTSR),[13,14] chosen because of its previous use in science disciplines and ease of assessing the construct.[11] Most of the groups contained one "high" reasoner, two "middle" reasoners, and one "low" reasoner. The designations of high, medium, and low are norm-referenced groupings; the students were ranked from lowest score to highest score and then placed in groups. The activities are built on a learning cycle framework, which engages students in a cycle of exploring new or unfamiliar phenomena (exploration), testing their ideas against what is known (concept introduction), and formulating new hypotheses about the phenomena and applying them to future scenarios (concept application).[15] These interventions are also student-centered; the instructor circulates to answer questions and clarify points of confusion. This curricular design and teaching method was selected because it has been shown to encourage the development of SR skills.[4,16,17] The interventions provide students with data (in graphs or tables) or a short passage to read and then pose questions that require a particular scientific reasoning skill to arrive at an answer. To keep students on pace throughout the class period, content checkpoints via a student response system (clickers) were used at 10–15 min intervals. An example of checkpoints associated with Intervention 4, the Chernobyl disaster, can be found in Figure 1.

From the introductory paragraph, what caused the accident?

1) An explosion
2) An interruption of coolant during a routine test
3) An insufficient number of control rods in the reactor
4) 2 and 3

From question 10, how many half-lives of Cs-137 will it take to reach 1 mSv/year?

1) 50,000
2) 1
3) 16
4) 137

Figure 1. Example clicker checkpoints used with Intervention 4: Radioactive half-lives and Chernobyl.

In addition to the five interventions listed in Table 1, ten shorter activities were developed with the purpose of increasing the amount of student-centered activity and social discourse during the class period, which have also been suggested to help increase SR skills.[18] The shorter activities were designed to take approximately 10–15 min each and were focused on application of content or practice of skills without a deliberate emphasis on scientific reasoning skills or the learning cycle framework. The content of the shorter activities is shown in Table 2. These shorter activities were implemented at the end of the lesson preceding a full-class intervention and at the end of the class following the full-class intervention. Unlike the longer interventions, these short activities did not introduce new content.

Table 2. Short Activity Content Aligned with Interventions

Intervention	Activity	Content Focus
1	1	The London Fog and Air Pollution[a]
	2	Classifying Matter (Elements, Compounds, Mixtures)
2	3	Drawing Lewis Structures
	4	Are High SPF Sunscreens Beneficial?[a]
3	5	Molecular Vibrations and IR Absorption[a]
	6	VSEPR Theory/Molecular Shapes
4	7	Balancing Nuclear Decay Equations
	8	Annual Radiation Dosage[a]
5	9	"Food Rules" Evaluation Activity
	10	Fatty Acids Structure Identification

[a]Activities that provide prerequisite knowledge, context, or further application of intervention topics.

Not all of the shorter activities that precede and follow the longer interventions have curricular connections to those interventions but rather simply address the topics that come before or after the content within the structure of the course. The activities that do have connections—providing prerequisite knowledge, context, or further application—are denoted with an "*a*" in Table 2 and are discussed with their respective interventions in the Supporting Information.

The design process began with deciding how many interventions to have across the semester and what content to address. The decision was made to design one intervention per textbook unit and to focus on content that would be easy enough for the students to work through and understand on their own but difficult enough to provide some challenge to those who may have been more familiar with the topic. Once the content foci of the interventions were determined, we examined the content in light of each of the reasoning skills assessed by the LCTSR: conservation of mass and volume; proportional, probabilistic, correlational, and hypothetico-deductive reasoning; and the controlling of variables. The goal was to identify which skills naturally aligned with the content, for example, addressing proportional reasoning while presenting material on the properties of light. Framing activities around the content first and reasoning skills second limited the scope of reasoning skills that could be addressed through the interventions, as not all skills (e.g., probabilistic reasoning, controlling of variables) fit with the course content. Drafts of the interventions underwent multiple rounds of revision and were pilot tested with a group of chemistry education research graduate students and faculty to determine if any questions were unclear. The pilot testers provided suggestions to clarify

dx.doi.org/10.1021/ed500207t | J. Chem. Educ. XXXX, XXX, XXX–XXX

Table 3. Average RTOP Subscale and Total Score by Class Type[a]

	Lect. 1	Lect. 2	Lect. 3	Lect. 4	Lect. 5	Lect. 6	Lect. 7	Lect. Median
LDI	4.3	5.0	4.7	5.7	4.7	3.7	4.7	4.7
C	17.7	15.3	14.3	17.3	14.7	14.3	16.3	15.3
CC	9.3	9.0	10.3	9.3	9.0	7.3	8.0	9.3
Total	31.3	29.3	29.3	32.3	28.3	25.3	29.0	29.3

	Act. 1	Act. 2	Act. 3	Act. Median
LDI	5.3	5.7	8.7	5.7
C	16.3	16.7	22.3	16.7
CC	15.0	14.3	17.3	15.0
Total	36.7	36.7	48.3	36.7

	Inter.1	Inter. 2	Inter.3	Inter.4	Inter. 5	Inter. Median
LDI	11.3	11.7	11.3	12.0	11.3	11.3
C	25.0	24.0	25.7	27.3	29.3	25.7
CC	25.3	22.0	26.7	21.7	23.3	23.3
Total	61.7	57.0	63.7	60.7	64.0	61.7

[a]LDI = Lesson Design and Implementation; C = Content; CC = Classroom Culture.

the wording of the instructions and questions. On the basis of those suggestions, revisions were made and were tested again with an undergraduate chemistry education researcher. His answers to questions in the interventions were taken into consideration for final revisions to arrive at the final products. Detailed summaries of the individual interventions and how specific questions focus on reasoning skills can be found in the Supporting Information.

In terms of logistics, the interventions were designed to take most, if not all, of a 50 min lecture period; getting students organized into their groups as they walk in is helpful. The activities were placed at the end of class periods that were more lecture-based; around 5 min was allotted for group organization. These procedures were in place from the beginning of the semester so students could see that group work was part of the class routine. It took little time to transition between group work and lecture.

■ EVALUATION

The evaluation of the intervention design and implementation was done to ensure that three goals were met: (1) the interventions were student-centered, (2) they were able to increase students' understanding of the course content, and (3) the interventions helped students increase their SR skills. The research tools used to evaluate each of the three goals are described in the subsections that follow. Prior to collecting any quantitative or qualitative data associated with the study, IRB approval was obtained from the university.

Student-Centeredness: Classroom Observations

To understand how the interventions and activities were functioning in the classroom environment, we observed and video recorded group interactions and listened to interactions between the instructor (second author) and student groups. In response to the observations, changes were made to the wording of some instructions and questions. No major changes were made that affected the content addressed in any of the interventions.

To investigate the differences between the instructor's normal lecture periods, class periods that included an activity, and class periods that consisted of only an intervention, video data of the three types of classes were scored using the Reformed Teaching Observation Protocol (RTOP).[19] The RTOP is a 25-item classroom observation protocol that is standards-based, inquiry-oriented, and student-centered and is designed to evaluate teaching practices that emphasize a problem-solving approach and active learning.[16] Each item on the RTOP is scored on a Likert scale from 0 (never occurred) to 4 (very descriptive) depending on how descriptive each statement was of the class period observed. While this makes the theoretical range of possible scores 0 to 100, extensive use of the instrument in our research group has shown that it has a practical lower limit of about 20. The RTOP has subscales of lesson design and implementation (LDI), content (C), and classroom culture (CC) in order to give instructors more pointed feedback on what can be improved in their classroom to make it more "reformed." To encourage high inter-rater agreement of the measurements, we used an improved RTOP instrument as described in a recent study with the RTOP.[20]

Over the course of the semester, 15 class periods were observed and video recorded: seven traditional lecture periods, three activity periods, and all five intervention periods. The videos were each scored by three trained raters, the first author being a rater on all 15 videos. If scores from independent scoring differed by more than five points, the scores were negotiated to reach score convergence within five points. For the 15 class periods observed, the average rating for each of the three subscales and the total score are presented in Table 3.

There are markedly different median scores for the various types of lessons observed throughout the semester. Higher scores tend to be associated with the interventions and lower scores with lectures. For further confirmation of this difference in instruction types, a Kruskal–Wallis analysis of the total RTOP score by class type showed a significant difference, $H(2) = 12.04$, $p = 0.002$. Since a significant difference existed, *post hoc* analyses were conducted to determine where the difference existed using three Mann–Whitney U tests. Because of the need for multiple comparisons, a Bonferroni correction was made to minimize the familywise error rate, making the cutoff for rejecting the null hypothesis $p = 0.0167$ for these analyses. When comparing the interventions to the activities, there is no significant difference between the median RTOP scores, $U < 0.001$, $p = 0.024$, $r = 0.75$. When comparing the interventions RTOP scores to those of the lectures, there is a significant difference, $U < 0.001$, $p = 0.003$, $r = 0.82$. There is also a significant difference when comparing the activities' RTOP scores to those of the lectures, $U < 0.001$, $p = 0.016$, $r = 0.76$. It is not surprising that no difference was detected between the

dx.doi.org/10.1021/ed500207t | J. Chem. Educ. XXXX, XXX, XXX−XXX

RTOP scores for the activities and those for the interventions, as there was not adequate statistical power due to fewer scores for both of these lesson types. The way in which the groups are coded in the statistical software package (SPSS Version 21), however, allows us to conduct a Jonckheere-Terpstra analysis to investigate if there is an increasing or decreasing trend to the medians of the groups being compared.[21] In other words, by coding the interventions as "1," the activities as "2," and the lectures as "3," we can use the Jonckherre test to see if the medians of the groups' total scores decrease, as would be expected. The results of the Jonckheere analysis confirm that as the amount of intended student-centered behavior decreased, the medians of the total RTOP scores in each group also decreased, $J < 0.001$, $z = -3.84$, $r = -0.99$. Both analyses were also performed on each of the subscales (LDI, C, CC) and produced similar results. These results demonstrate that the interventions and short activities that were developed are structurally different than this instructor's normal lecture style, which was one of our intended goals. The interventions and activities were designed to have students work in groups, with the teacher acting as a resource while students interact with models, graphs, and/or manipulatives related to the content addressed in the intervention or activity for that day. These improvements over a standard lecture setting resulted in higher scores on the RTOP items corresponding with students' use of models and the teacher serving as a resource. This confirmed and validated our design choices with observational data. Specifically looking at the Classroom Culture (CC) subscale scores for the different lesson types, it is apparent that the interventions scored well above the lectures with the activities falling in between, thus demonstrating that the activities and interventions, designed to be more student-centered, were executed as such. Examples of RTOP items in the CC subscale included "there was a high proportion of student talk and a significant amount of it occurred between and among students" and "the metaphor of 'teacher as listener' was very characteristic of this classroom."[19]

Student Content Understanding: Content Questions

To investigate students' content mastery gains from their use of the interventions and activities, exam items were written or used from previous semesters to assess content that was exclusively addressed by the interventions. Example exam questions related to the content addressed in the interventions can be found in Figure 2. Across four midterm exams (approximately 120 items in total), a total of 30 items were identified that evaluated content from the interventions and activities from the 83 students in the intervention semester.

1) Using the **adjusted line in the ozone graph** for Oxford (See Figure 2), what is the approximate rate of change **per year** from 2001 to 2004?
A. 4 ppb/year B. 0 ppb/year C. 3 ppb/year
D. 7 ppb/year E. 1 ppb/year

2) Of the radiant energies (E) listed, which would have the **longest wavelength**?
A. 1.0×10^{-15} Joules B. 1.0×10^{-25} Joules
C. 4.0×10^{-8} Joules D. 9.0×10^{-15} Joules
E. 9.0×10^{-8} Joules

Figure 2. Example exam questions related to intervention content.

We hoped that students would perform well on these items; however, it was difficult to examine the results without a basis for comparison. Consequently, we identified 26 items that assessed chemistry content aligned with the interventions from the four midterm exams of a previous semester (180 students) to serve as a comparison. Data for student performance on these items in both semesters are shown in Figure 3.

From the data presented in Figure 3, it appears that students in the intervention implementation semester performed better on average than those in the semester without the interventions. To determine whether the difference between semesters was significant, a Mann−Whitney U test was conducted, as the assumptions of normality and homogeneity of variances for an ANOVA were violated.

The analysis revealed that there is a significant difference between the two groups on their aggregated scores across the midterm exams with a small effect size, $U = 5601.0$, $z = 3.26$, $p = 0.001$, $r = 0.20$. At this point, it is important to note a limitation of this comparison. Although the items from both semesters assess the same content, they do not use the same wording from semester to semester, which could cause students to respond or think about the questions in different ways. To evaluate the extent of this limitation, we calculated the item difficulty for each item in each semester. The comparison semester had a range of item difficulties from 0.33 to 0.95 with an average difficulty of 0.69. The intervention semester had a range of item difficulties from 0.34 to 0.96 with an average difficulty of 0.72. These results suggest that over the course of the semester, the two groups of students performed similarly. Even with the previously discussed limitation, it is a promising result that students are able to answer questions correctly about content that they investigated through student-centered interventions to a slightly higher degree than those that did not have exposure to the interventions. This finding provides further evidence that reformed teaching practices that actively engage students in a more active role result in positive knowledge outcomes. This result is not surprising, however, given the body of literature that suggests use of student-centered teaching practices framed around a constructivist[3−5,15,22] learning theory.

Student SR Ability: Analogous Reasoning Task Interviews

After each of the first four interventions, one-on-one, semistructured interviews were conducted with a small number of students to verify both the content gains of the activity (by asking the participants what they learned through the interventions) and the reasoning gains (by asking the participants to answer a question analogous to those that appear on the LCTSR).[13,14] No students participated in more than one interview throughout the semester. The interviews were conducted within a week of each intervention to try to ensure that the growth in their content and reasoning could be attributed to the activities, not other factors in the classroom environment or maturation. The results from the reasoning portion of the interview are presented in Table 4a−d. The answer profile column shows the possible combinations of performance on the LCTSR pretest and on the analogous reasoning task in the interview setting. The frequency column shows the number of students who exhibited that particular profile from that set of interviews. The analogous questions came from a second measure of scientific reasoning, the Test of Logical Thinking (TOLT).[23] Table 4a,b,d shows that there is a fair number of students in each set of interviews who improve following the intervention. After Intervention 1, three out of the eight students who completed the interview answered both pretest items incorrectly but were able to correctly answer the

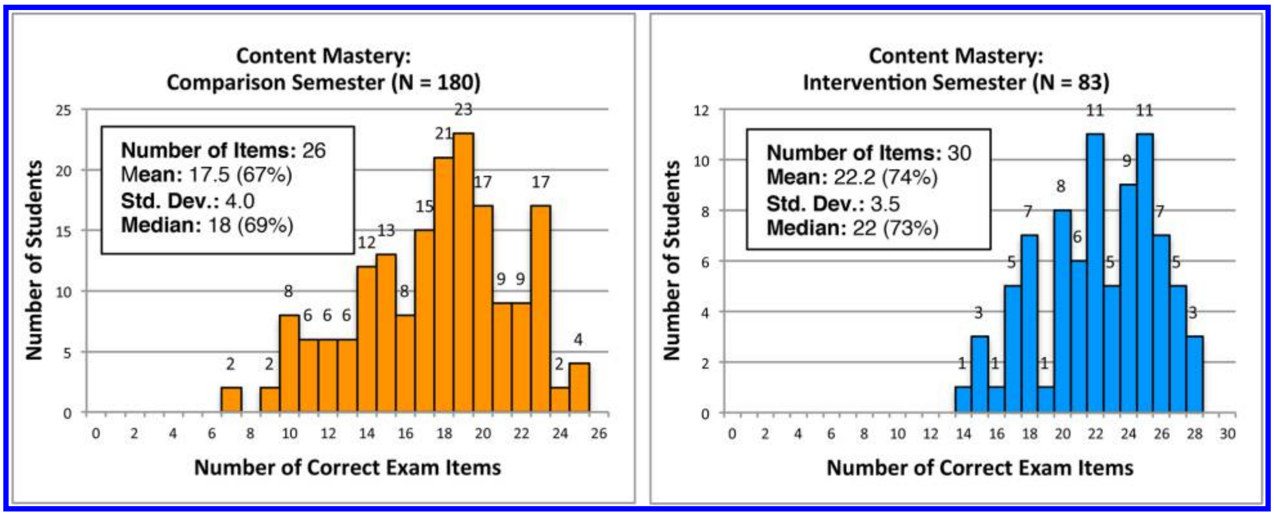


Figure 3. Content mastery data for comparison and intervention semesters.

Table 4. Frequency of Answer Profiles on LCTSR (pretest) and Analogous (interview) Items

Answer Profile (pretest → interview)	Frequency
(a) Intervention 1 (proportional, 8 participants)	
Both incorrect → correct	3
One incorrect → correct	0
One incorrect → incorrect	1
Both correct → correct	4
Both incorrect → incorrect	0
(b) Intervention 2 (proportional, 7 participants)	
Both incorrect → correct	3
One incorrect → correct	1
One incorrect → incorrect	1
Both correct → correct	1
Both incorrect → incorrect	1
(c) Intervention 3 (correlational, 5 participants)	
Incorrect → correct	0
Incorrect → incorrect	2
Correct → correct	0
Correct → incorrect	3
(d) Intervention 4 (proportional, 7 participants)	
Both incorrect → correct	3
One incorrect → correct	1
One incorrect → incorrect	1
Both correct → correct	1
Both incorrect → incorrect	1

analogous item in the interview. After Intervention 2, three out of the seven students who completed the interview had both pretest items incorrect and answered the analogous item correctly; another three of seven performed similarly after Intervention 4.

The negative results of the interviews after Intervention 3 (Table 4c) suggest that the reasoning questions in the intervention likely did not effectively map onto the reasoning skill tested through the analogous question. Alternatively, one could conclude that students "lost" reasoning skills as a result of the intervention; however, a student regressing to more naïve reasoning patterns does not align with the Piagetian developmental framework underlying SR skills.

Looking at the results across all three proportional reasoning interventions (Interventions 1, 2, and 4), 6 of the 22 students consistently answered correctly on the pretest and interview items. Of the remaining 16, 9 students showed growth from the pretest to the interview. Of the seven remaining, 2 improved and 5 did not. It should also be noted that students that demonstrated growth in their answer profile (both incorrect → correct) across all of the interviews had a range of pretest scores from four to eight, suggesting that the interventions help students improve or maintain their skills across levels of overall SR ability.

■ CONCLUSIONS

On the basis of the evaluation outcomes, we assert that the interventions met their intended goals. First, the amount of student-centeredness was apparent through use of the RTOP, which not only shows that these materials are significantly different than a lecture period but also that the interventions were being implemented by the instructor with fidelity. Second, students using the interventions were able to gain as much or more content knowledge than those who did not. The interventions were also effective at improving the reasoning skills in some students, as evident in their interview performance as compared to their performance on the pretest measure. The evaluation results validated that the multiple revisions before implementation helped to produce activities aligned with the goals of the reform.

Secondary to the major findings is the notion that classroom activities should be subjected to the detailed and rigorous evaluation that various research tools offer to enable our students to interact with the highest quality materials possible. This is especially important with the population at hand (nonscience majors), as the course may be the only exposure some students have to college-level chemistry instruction. We hope that instructors teaching nonmajors chemistry will find the interventions to be useful teaching resources and that our evaluation approach contributes to the development of future classroom materials.

■ FUTURE WORK

The interventions described have been revised since these data were collected, and the documents found in the Supporting Information are the most current versions. The intervention implementation took place over the course of a second semester with the revised materials. The data from this group of students will be analyzed in concert with that of the initial implementation group and the semesters without implementation to determine gains in SR attributed to the use of these targeted materials in the classroom.

ASSOCIATED CONTENT

Supporting Information

Descriptions and full versions of all interventions can be found in the Supporting Information. This material is available via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: yeziere@miamioh.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank all of the students who participated in the study, as well as the Yezierski and Bretz research groups at Miami University for their invaluable input and suggestions.

REFERENCES

(1) American Association of Colleges & Universities. http://www.aacu.org/resources/liberaleducation/index.cfm (accessed Feb. 2014).
(2) Foundation for Critical Thinking. http://www.criticalthinking.org/pages/defining-critical-thinking/766. (accessed Feb. 2014).
(3) White, B.; Stains, M.; Escriu-Sune, M.; Medaglia, E.; Rostamnjad, L.; Chinn, C.; Sevian, H. A Novel Instrument for Assessing Students' Critical Thinking Abilities. *J. Coll. Sci. Teach.* **2011**, *40* (5), 102−107.
(4) Benford, R.; Lawson, A. E. *Relationships Between Effective Inquiry Use and the Development of Scientifc Reasoning Skills in College Biology Labs*; No. ED456157; ERIC Document Reproduction Service: 2001.
(5) Marshall, J. C.; Horton, R. M. The Relationship of Teacher-Facilitated Inquiry-Based Instruction to Student Higher Order Thinking. *School Sci. Math.* **2011**, *111* (3), 93−101.
(6) Farrell, J. J.; Moog, R. S.; Spencer, J. M. A Guided Inquiry General Chemistry Course. *J. Chem. Educ.* **1999**, *76* (4), 570−574.
(7) Moog, R. S.; Farrell, J. J. The POGIL Project. *Chemistry: A Guided Inquiry*, 6th ed.; Wiley: New York, 2014.
(8) Lantz, J.; Cole, R. The POGIL Project. *Analytical Chemistry: A Guided Inquiry Approach Quantitative Analysis Collection*; Wiley: New York, 2013.
(9) Spencer, J. M.; Moog, R. S.; Farrell, J. J. The POGIL Project. *Physical Chemistry: A Guided Inquiry - Thermodynamics*; Wiley: New York, 2012.
(10) Spencer, J. M.; Moog, R. S.; Farrell, J. J.; The POGIL Project. *Physical Chemistry: A Guided Inquiry − Atoms, Molecules, and Spectroscopy*; Wiley: New York, 2012.
(11) Carmel, J. H.; Yezierski, E. J. Are We Keeping the Promise? Investigation of Students' Critical Thinking Growth. *J. Coll. Sci. Teach.* **2013**, *42* (5), 71−81.
(12) American Chemical Society. *Chemistry in Context: Applying Chemistry to Society*, 7th ed.; McGraw-Hill: New York, 2009.
(13) Lawson, A. E. The Development and Validation of a Classroom Test of Formal Reasoning. *J. Res. Sci. Teach.* **1978**, *15*, 11−24.
(14) Lawson, A. E. *Classroom Test of Scientific Reasoning: Revised*; Arizona State University: Tempe, AZ, 2000.
(15) Karplus, R. Teaching for the Development of Reasoning. *Res. Sci. Educ.* **1980**, *10*, 1−10.
(16) Lawson, A. E.; Wollman, W. T. Encouraging the Transition from Concrete to Formal Cognitive Functioning − An Experiment. *J. Res. Sci. Teach.* **1976**, *13* (5), 413−430.
(17) Wollman, W. T.; Lawson, A. E. The Influence of Instruction on Proportional Reasoning in Seventh Graders. *J. Res. Sci. Teach.* **1978**, *15* (3), 227−232.
(18) Gerber, B. L.; Cavallo, A. M. L.; Marek, E. A. Relationships Among Informal Learning Environments, Teaching Procedures and Scientific Reasoning Ability. *Int. J. Sci. Educ.* **2001**, *23* (5), 535−549.
(19) Sawada, D.; Piburn, M. D.; Judson, E.; Turley, J.; Falconer, K.; Benford, R.; Bloom, I. Measuring Reform Practices in Science and Mathematics Classrooms: The Reformed Teaching Observation Protocol. *School Sci. Math.* **2002**, *102* (6), 245−253.
(20) Yezierski, E. J.; Herrington, D. G. Improving Practice with Target Inquiry: High School Chemistry Teacher Professional Development that Works. *Chem. Educ. Res. Prac.* **2011**, *12*, 344−354.
(21) Field, A. *Discovering Statistics Using SPSS (3rd ed.)*; Sage Publications: Thousand Oaks, CA, 2009.
(22) Lawson, A. E. A Review of Research on Formal Reasoning and Science Teaching. *J. Res. Sci. Teach.* **1985**, *22* (7), 569−617.
(23) Tobin, K. G.; Capie, W. The Development and Validation of a Group Test of Logical Thinking. *Educ. Psych. Meas.* **1981**, *41*, 413−423.