JOURNAL OF CHEMICAL EDUCATION

Article

pubs.acs.org/jchemeduc

Preparation for College General Chemistry: More than Just a Matter of Content Knowledge Acquisition

Mark S. Cracolice* and Brittany D. Busby

Department of Chemistry and Biochemistry, University of Montana, Missoula, Montana 59812, United States

Supporting Information

ABSTRACT: This study investigates the potential of five factors that may be predictive of success in college general chemistry courses: prior knowledge of common alternate conceptions, intelligence, scientific reasoning ability, proportional reasoning ability, and attitude toward chemistry. We found that both prior knowledge and scientific reasoning ability were significantly correlated with students' performance on the American Chemical Society Division of Chemical Education Examinations Institute First Term General Chemistry Examination. Given that scientific reasoning ability was significantly correlated with final exam performance and that its impact is not broadly known in the chemistry teaching community, we then discuss the implications for facilitating the development of reasoning ability in college preparatory high school chemistry courses and college preparatory chemistry courses.

KEYWORDS: *High School/Introductory Chemistry, First-Year Undergraduate/General, Chemical Education Research, Curriculum, Testing/Assessment, Constructivism, Learning Theories*

FEATURE: Chemical Education Research

INTRODUCTION

It is no secret to readers of this *Journal* that college general chemistry is perceived by students as a relatively difficult course.[1] The perception is exacerbated by the fact that the majority of students in general chemistry courses are neither chemistry nor biochemistry majors.[2] Evidence from actual measurement of failure rates supports students' perceptions: a recent meta-analysis of 225 published and unpublished research studies in undergraduate science, technology, engineering, and mathematics (STEM) courses reported that the average failure rate (defined as being issued a D or F grade or withdrawing from the course) in STEM courses taught via traditional lecturing was 33.4%.[3] In general chemistry specifically, selected institutional historical failure rates of over 50% have been reported.[4,5]

Performance in general chemistry is not just a function of what happens in the course itself; we must look beyond issues related to course content and pedagogy. A correlate of performance is the state of knowledge of a student at the beginning of a course, both declarative (knowing *that*) and procedural (knowing *how*). Students enter the general chemistry course neither with their minds completely devoid of knowledge of the content of general chemistry nor absent of some of the procedural knowledge needed to succeed (see, e.g., refs 6−11). The importance of the declarative knowledge portion of the continuum has been recognized by chemical educators for about 100 years.[12] Accordingly, the American Chemical Society Division of Chemical Education Examinations Institute (abbreviated as ACS-EI hereafter) currently offers two undergraduate placement exams designed to assess the readiness of students for the general chemistry course.[13] The California Chemistry Diagnostic Test is a mixture of questions of chemistry word problems, chemistry concepts, and math word problems.[14] The Toledo Examination is divided into three parts: mathematics, general chemistry knowledge, and specific—primarily descriptive—chemistry knowledge.[15,16] It is important to note, in particular, that both exams are primarily declarative-knowledge-based.

Placement exams are used to place students into either the general chemistry sequence or a preparatory chemistry course. Students who prepared for college general chemistry via a high school chemistry course would have to retain sufficient declarative knowledge to be able to obtain the institutional cut score on a placement exam. With both high school chemistry for college-bound students and college preparatory chemistry, most courses of either of these types focus on preparing students for general chemistry by teaching what is effectively a simplified general chemistry course. Development of students' scientific reasoning ability does not appear to be the focus of instruction at any level as it is not typically included in the standards or the assessments. For example, the *Science Content Standards for California Public Schools: Kindergarten through Grade Twelve* requires a high school chemistry course to provide students with the opportunity to learn the following topics:[17]

(1) atomic and molecular structure
(2) chemical bonds
(3) conservation of matter and stoichiometry
(4) gases and their properties
(5) acids and bases
(6) solutions
(7) chemical thermodynamics
(8) reaction rates
(9) chemical equilibrium
(10) organic chemistry and biochemistry

(11) nuclear processes

Note that every topic in the California high school chemistry curriculum is based on unobservable causal entities and thus requires fully developed scientific reasoning skills for students to be able to construct their declarative knowledge.

Collectively, the methods that the chemical education community currently uses to establish widely followed standards for high school chemistry courses and assess readiness for college general chemistry are based on declarative knowledge. However, the high average failure rate for STEM courses indicates that something may be amiss with these methods. This study therefore is designed to serve as an early, basic exploration into what may be missing in preparing students for success in college general chemistry.

Investigating Factors in Addition to Declarative Knowledge

Research has indicated that factors other than initial declarative knowledge may be predictive of performance in general chemistry. Intelligence is perhaps the most well-known cognitive characteristic that has research evidence to indicate that it is predictive of success in the classroom. A task force convened by the American Psychological Association defined intelligence by stating, "Individuals differ from one another in their ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought. Although these individual differences can be substantial, they are never entirely consistent: A given person's intellectual performance will vary on different occasions, in different domains, as judged by different criteria. Concepts of "intelligence" are attempts to clarify and organize this complex set of phenomena."[18]

Knowledge of topics that commonly are associated with alternate conceptions is another potential predictor of success in general chemistry. An alternate conception is a mental construct held by an individual about the natural world that is different from the construct generally agreed upon by the scientific community.[6,19−21] A diverse set of alternate conceptions are held by students entering a general chemistry course, and they are resistant to being modified by conventional instruction.[19] Students hold alternate conceptions about topics typically taught in first-semester general chemistry such as understanding what a balanced chemical equation represents, the mole concept, and thinking about atoms and molecules in general.[21]

Scientific reasoning ability also has been demonstrated to predict success in general chemistry. Scientific reasoning ability is "a general disposition to consider alternative possibilities and the acquisition of accompanying hypothesis testing schemes that allow one to process evidence to choose among the alternatives (e.g., control of variables, correlational reasoning, probabilistic reasoning)."[22] Lewis and Lewis[23] found that scientific reasoning can be used to predict students at risk in general chemistry: students with poor scientific reasoning ability are likely to have low ACS-EI examination scores. Scientific reasoning ability also has been shown to be related to the ability to solve conceptual chemistry problems successfully.[24−26]

Proportional reasoning is one type of reasoning within the set of skills that collectively comprise scientific reasoning ability. Proportional reasoning can be modeled as being composed of (1) analogical reasoning, including the ability to solve verbal analogies and numerical analogies, (2) routine proportionality, which is the ability to solve proportional tasks, and (3) meta-analogical awareness, which consists of the ability to recognize proportional and nonproportional statements and the ability to work with nonproportional tasks.[27] Given that the standard first-semester general chemistry curriculum has multiple topics based on proportional reasoning, for example, unit conversions, gas laws, stoichiometric relationships, and solution concentration, it is reasonable to hypothesize that proportional reasoning ability may influence student content knowledge upon completion of the course.

Attitude is another potential predictor of success in general chemistry. Chan and Bauer[28] found that attitude and self-concept were collectively predictive of exam performance in first-semester general chemistry. They proposed that chemistry achievement is predicated upon cognitive characteristics, affective experiences such as self-concept and attitude, and other motivational and cognitive processes. Self-concept itself has been shown to positively correlate to performance on the First-Term General Chemistry ACS Exam.[29]

Theory Base and Research Questions

There is more to preparing students to succeed in general chemistry than simply facilitating the development of declarative knowledge. Development of students' *procedural knowledge* matters as well. The late Swiss psychologist Jean Piaget (1896−1980) originally discovered two critical concepts related to the development of procedural knowledge. First, the ability to utilize and apply scientific reasoning patterns reflectively begins to emerge with the onset of adolescence.[30−32] Second, students develop their scientific reasoning abilities only when in an environment that facilitates their development.[33−35] The background discussed to this point and the theory base led to the research question:

> *Are prior alternate conception knowledge, prior procedural knowledge (intelligence, scientific reasoning ability, and proportional reasoning ability), and entering attitude predictive of students' success on the First-Term General Chemistry ACS Exam as students complete a semester of instruction?*

■ METHODS

Ethics Statement

Permission for human participants use was obtained by the Institutional Review Board of the authors' university, and consent was obtained from all participants. None of the participants elected to opt out.

Course and Participants

The participants were enrolled in a first-semester general chemistry course at a 10,000-undergraduate-student-enrollment state university with a 96% admission acceptance rate and a 4-year graduation rate of 24%. The course was designed to provide students the opportunity to learn the standard U.S. first-semester content, as inferred through observation of the contents of the ACS-EI First-Term General Chemistry exams and the contents of the best-selling general chemistry textbooks over the past two decades. It was taught via a combination of pedagogies including adaptations of guided inquiry (see, e.g., refs 36 and 37), peer-led team learning (see, e.g., refs 38 and 39), and the flipped classroom (see, e.g., refs 40 and 41). The course was not modified in any way because of the research

DOI: 10.1021/acs.jchemed.5b00146
J. Chem. Educ. XXXX, XXX, XXX−XXX

project other than structuring time in the laboratory portion of the course for administration of the instruments.

Experimental Design

A correlational research design was employed.[42] Five pretests, described in the following, were administered in the first 2 weeks of the 16-week semester course during the regularly scheduled laboratory meeting. Ten total extra points (as compared with 1000 possible earned course points) were possible based on pretest participation. In the first week, knowledge of topics associated with alternate conceptions (2 points), scientific reasoning ability (2 points), and attitude toward chemistry (1 point) were measured. These characteristics were judged to have the most potential to change based on instruction. In the second week, intelligence (2 points) and proportional reasoning ability (2 points) were measured, and demographic data (1 point) was collected. The time allotted for administration of each instrument was based on experience with administering the instruments in the previous academic year, and it was sufficient for what we observed as allowing every student to finish completely and without time pressure. A graduate student in chemistry education research and article coauthor (B.D.B.) supervised the administration of all pretests. To encourage the participants' best effort, they were informed that the full time allotted for each instrument would be fully expired before beginning the next instrument or the scheduled laboratory exercise. The post-test, also described in the following, was administered during the scheduled final examination period in the 16th week of the semester.

Instruments

Knowledge of Topics Associated with Alternate Conceptions: Chemistry Concepts Inventory (CCI). The Chemistry Concepts Inventory (CCI) is a 16-item multiple-choice instrument designed to measure the extent of alternate conceptions about topics typically covered in a first-semester general chemistry course.[43] It consists of 22 multiple-choice questions, with 10 stand-alone questions and six paired-question sets. We scored the instrument requiring both of the questions in a paired-question set to be correct for credit. Participants were given 20 min to complete the instrument. Barbera[44] conducted a psychometric analysis of the instrument and concluded that it is suitable for large-scale assessment of student understanding. The instrument is both reliable (Cronbach's α ranged between 0.704 and 0.76 in various trials) and reproducible (test-retest correlation = 0.79). The validity of the instrument was initially established by a panel of experienced chemical education researchers, and it has been used in many published, peer-reviewed research studies (see, e.g., refs 45 and 46).

Intelligence: Raven Standard Progressive Matrices—Plus Version (SPM+). Intelligence was measured with the Raven Standard Progressive Matrices—Plus Version (SPM+); the Plus Version was developed to restore discriminative ability at the upper limit because a ceiling effect began to appear with the original version.[47] The instrument consists of five sets of 12 matrices of increasing difficulty, for a maximum score of 60. Participants were given 60 min to complete the instrument. The validity of the instrument has been established over years of use; it is widely accepted as one of the best large-group measures of nonverbal intelligence. Many forms of reliability data have been collected; its internal consistency reliability was measured as 0.88, its test-retest reliability is 0.97, and its criterion-related validity has been established through numerous studies that indicate a relationship between the SPM score and job performance.[48]

Scientific Reasoning Ability: Classroom Test of Scientific Reasoning (CTSR). The Classroom Test of Scientific Reasoning, Multiple Choice Version (CTSR), is a 12-item multiple-choice instrument designed to measure scientific reasoning ability.[49] It consists of 24 paired-questions sets, where a student must answer both items in a pair correctly for credit. Participants were allowed 35 min to complete the instrument. Validity was originally established by a panel of experts in Piagetian research, and the instrument has been used in many published, peer-reviewed research studies (see, e.g., refs 33 and 50). A subsequent study by Lawson[51] found the instrument to be reliable (Cronbach's α = 0.70).

Proportional Reasoning Ability: Paper and Pencil Balance Beam Task (PPBBT). The Paper and Pencil Balance Beam Task (PPBBT) is a 36-item instrument designed to measure proportional reasoning ability.[52] Participants were allowed up to 30 min to complete the instrument. Concurrent validity was established by comparison of scores on the paper-and-pencil version and clinical interviews; the paper-and-pencil task was based on Inhelder and Piaget's[30] "equilibrium in the balance" task. The scores were nearly identical. Construct validity evidence was established by comparing performance of sixth- and ninth-graders, finding that the ninth-graders were more likely to use more sophisticated strategies. A separate study of construct validity also showed that age was significantly correlated with number of correct responses in a group of 8−15 year-olds. Test-retest reliability was correlated at 0.85 over a two-week interval, and very similar results were obtained with alternate forms of the instrument.[52] It is also notable that the PPBBT was initially piloted with a 21-participant sample of adult university students, and 48% of the adult sample incorrectly considered only one variable (weight *or* distance from the fulcrum) in their responses as compared to 46% of the 8−15-year-old participants.[52]

Attitude toward Chemistry: Attitude toward the Subject of Chemistry Inventory (Version 2) (ASCI V2). Attitude toward chemistry was measured with the Attitude toward the Subject of Chemistry Inventory (Version 2) (ASCI V2) instrument that was originally developed by Bauer[53] and subsequently modified by Xu and Lewis[54] to improve its construct validity and decrease the time needed for administration. It consists of eight pairs of adjectives that are divided into two subscales, intellectual accessibility and emotional satisfaction. Participants were given 3 min to complete the instrument. Xu and Lewis[54] provided data illustrating the construct validity of the instrument via confirmatory factor analysis, which supported a two-factor model with a comparative fixed index value of 0.95. Cronbach's α was measured as 0.82 for the intellectual accessibility subscale and 0.79 for the emotional satisfaction subscale.

Post Content Knowledge: ACS-EI First Term General Chemistry Examination, Form 2009 (ACS). The ACS-EI First Term General Chemistry Examination is a 70-item instrument designed to measure general chemistry content knowledge upon completion of one semester of the sequence.[55] Participants were allowed 120 min to complete the instrument, which was used as the final examination for the course. It constituted 30−50% of a student's grade, contingent upon performance on the midterm exams. Construct, face, and content validity are based on the fact that the exam was developed by a 15-person panel of general chemistry experts

DOI: 10.1021/acs.jchemed.5b00146
J. Chem. Educ. XXXX, XXX, XXX−XXX

and then field tested and refined by the panel. Norms are based on the scores of 3827 students in 34 colleges. The Kuder-Richardson KR-21 coefficient of reliability is 0.88. The norm mean score is 37 (SD = 11), and the norm median score is 36.[56]

■ RESULTS AND DISCUSSION

Of the 157 participants who took the final exam, there were complete data sets for 135, and only the complete sets were used in the analysis. Table 1 summarizes the abbreviations used

Table 1. Abbreviations Used in Tables

Abbreviation	Instrument	Characteristic Measured
CCI	Chemistry Concepts Inventory	Alternate conceptions about topics typically covered in first-semester general chemistry
CTSR	Classroom Test of Scientific Reasoning	Scientific reasoning ability
SPM+	Raven Standard Progressive Matrices—Plus Version	Intelligence
ASCI V2	Attitude toward the Subject of Chemistry Inventory (Version 2)	Attitude toward chemistry
PPBBT	The Paper and Pencil Balance Beam Task	Proportional reasoning ability
ACS	ACS-EI First Term General Chemistry Examination	General chemistry content knowledge

for the six instruments. The descriptive results are summarized in Table 2. Distribution plots and tables for each instrument are available in the Supporting Information. All distributions were reasonably normal.

Table 3 illustrates rough measures of internal reliability for this particular experiment. It compares the statistics from four of the five instruments administered as a pilot pretest in the previous academic year with the statistics obtained for the academic year in which the study was conducted. (The PPBBT is not included in this comparison because it was piloted as a post-test only in the previous academic year.) The institution at which the pilot and experiment were conducted made no change to admissions policy from one year to the next. Inspection of the data shows that there is high sample-to-sample reliability, producing similar results under constant conditions.

The Pearson product-moment correlation coefficient matrix for all instruments is provided in Table 4. We selected a significance level of $p < 0.05$ because there were no compelling reasons to deviate from the generally accepted standard.[57] Note that at the pairwise correlation-only level of analysis, four of the five pretest instruments plus the effect of the course have a statistically significant positive correlation with the post-test.

Analysis of variance statistics are given in Table 5 for the predictors (constant), CCI, CTSR, SPM+, ASCI V2, and

Table 3. Internal Reliability Statistics

Instrument	CCI		CTSR		SPM+		ASCI V2	
Pilot or Expt	Pilot	Expt	Pilot	Expt	Pilot	Expt	Pilot	Expt
Mean	7.0	7.1	7.7	7.7	44.8	43.9	8.0	7.9
SD	3.0	2.8	2.1	2.0	4.8	5.0	1.5	1.4
Median	7	7	8	8	45	44	8	8

Table 4. Pearson Product-Moment Correlation Coefficient Matrix for All Instruments

	CCI	CTSR	SPM+	ASCI V2	PPBBT	ACS
CCI	1.000	0.441[a]	0.257[a]	0.355[a]	0.169	0.448[a]
CTSR		1.000	0.404[a]	0.064	0.207[b]	0.478[a]
SPM+			1.000	0.114	0.199[b]	0.310[a]
ASCI V2				1.000	0.062	0.141
PPBBT					1.000	0.243[a]
ACS						1.000

[a]Correlation is significant at the p = 0.01 level (two-tailed). [b]Correlation is significant at the p = 0.05 level (two-tailed).

Table 5. Analysis of Variance Statistics

	Univariate Analysis of Variance			Between-Subjects Effects	
	F	p	R^2	Adjusted R^2	Partial Eta^2
CCI	3.052	0.001	0.247	0.166	0.247
CTSR	5.102	< 0.001	0.245	0.197	0.245
SPM+	1.303	0.177	0.230	0.053	0.230
ASCI V2	1.224	0.294	0.063	0.012	0.063
PPBBT	1.826	0.030	0.0221	0.100	0.221

Multiple Linear Regression			
F	p	R^2	Adjusted R^2
12.221	< 0.001	0.321	0.295

PPBBT, plus the effect of the course, on the dependent variable ACS. A statistically significant effect is indicated. The multiple regression results are summarized in Table 6. Since two or more variables may be measuring the same construct (based on the correlation coefficients in Table 4), we now see that we may conclude that the CCI and the CTSR, plus the effect of the course, are significantly correlated with the ACS with effect sizes of 0.251 and 0.296, respectively. By using Cohen's[58] guidelines for expressing effect sizes as small, medium, or large, the correlations of prior content understanding (r = 0.45) and scientific reasoning ability (r = 0.47) are classified as medium (between r = 0.30 and 0.50). The results show that both alternate conceptions about topics typically covered in first-semester general chemistry and scientific reasoning ability, as measured at the beginning of the course, influence general chemistry content knowledge after a semester of instruction. Figure 1 graphically illustrates the data for the pretest—post-test pairs that are correlated at a statistically significant level. Three of the pretests plus the effect of the course did not have

Table 2. Summary Statistics for Each Measurement Instrument

Instrument (Range)	CCI (0–16)	CTSR (0–12)	SPM+ (0–60)	ASCI V2 (0–16)	PPBBT (0–36)	ACS (0–70)
Mean	7.1	7.7	43.9	7.9	26.8	47.7
SD	2.8	2.0	5.0	1.4	3.7	12.1
Median	7	8	44	8	27	50
High	15	11	57	13	36	68
Low	2	3	31	5	16	10

DOI: 10.1021/acs.jchemed.5b00146
J. Chem. Educ. XXXX, XXX, XXX–XXX

Table 6. Multiple Regression Results for All Pretests with the ACS Post-test

Pretest	Coefficients	Standard Error	t Statistic	p-level	Null hypothesis (5%) rejected?	Cohen's f^2 Effect Size	95% Confidence Interval (Lower−Upper)
CCI	1.134	0.381	2.979	0.003	Yes	0.251	0.424−1.918
CTSR	1.787	0.523	3.418	< 0.001	Yes	0.296	0.738−2.813
SPM+	0.235	0.194	1.209	0.229	No	0.107	−0.148−0.622
ASCI V2	0.229	0.664	0.345	0.731	No	0.020	−1.209−1.333
PPBBT	0.390	0.249	1.563	0.120	No	0.063	−0.105−0.882
Intercept	3.347	10.225	0.327	0.744	No		−15.571−24.463

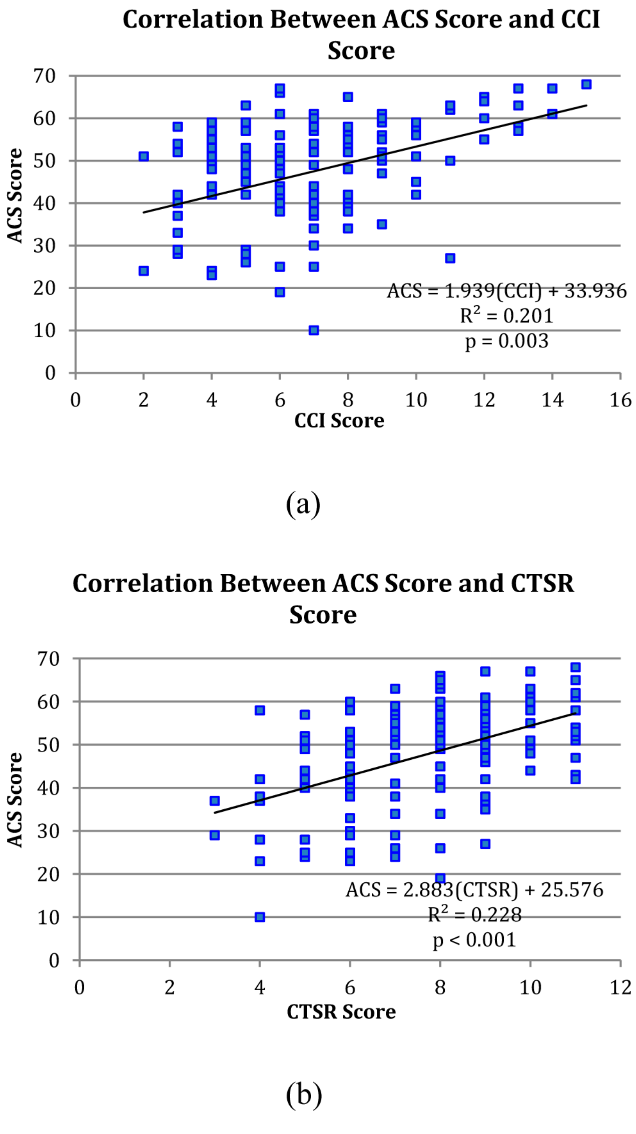

Figure 1. (a) Correlation between ACS score and CCI score. (b) Correlation between ACS score and CTSR score.

significant correlations with the post-test: intelligence, attitude toward chemistry, and proportional reasoning ability.

Addressing the research question:

Are prior alternate conception knowledge, prior procedural knowledge (intelligence, scientific reasoning ability, and proportional reasoning ability), and entering attitude predictive of students' success on the First-Term General Chemistry ACS Exam as students complete a semester of instruction?

Only prior alternate conception knowledge and scientific reasoning ability, as combined with the effect of the semester of instruction, were correlated with post content knowledge at a statistically significant level. They are correlated to students' success on the First-Term General Chemistry ACS Exam as students complete a semester of instruction.

Study Limitations

A correlational research design is not experimental—we cannot say that good prior content understanding and well-developed scientific reasoning ability *cause* high ACS exam scores; instead, these results say that there is a medium-strength relationship between two of the pretests and the effect of instruction with the post-test. It could be that the characteristics measured by the pretests caused the post-test result, what led to the post-test result caused the pretests' characteristics, the pretests' characteristics and the post-test result influenced one another, or some characteristic not measured in this study influenced performance on both the pretests and post-test. Additional research is needed to continue to explore this topic.

The study is also limited by generalizability concerns, that is, external validity. The course had a very unique curriculum design, and this design likely uniquely influenced the post-test results. Also, although the validity of four of the five pretest instruments is sufficiently established with a college-aged sample, the PPBBT was primarily validated with participants 15 years of age and younger, so it may or may not be a good measure of proportional reasoning ability as it is applied in first-semester general chemistry (however, note that the histogram in the Supporting Information shows a fairly normal distribution with our college-aged sample, and a small sample of university students in the original validation studies made the same errors as middle and high school students).

■ SUMMARY AND IMPLICATIONS FOR RESEARCH AND CHEMISTRY LEARNING

Even after a full semester of instruction in college general chemistry, students' prior alternate conception knowledge and their scientific reasoning ability at the onset of the course remain positively correlated at a medium effect size level to their performance on the ACS-EI First Term General Chemistry Examination administered at the end of the 16-week semester. The Chemistry Concepts Inventory and Classroom Test of Scientific Reasoning or other parallel instruments measuring the same constructs may be revelatory of the kind of knowledge that predicts outcomes in college general chemistry. Further research on the intersection of these cognitive constructs and instruction may be worth a good deal of additional study.

With respect to alternate conceptions, this study adds to the already substantial literature[20] that indicates that these misconceptions, naïve conceptions, preconceptions, private or personal models, or prescientific conceptions, as they are also known, are resistant to extinction by instruction. We believe that our sample is typical of students prepared for college

DOI: 10.1021/acs.jchemed.5b00146
J. Chem. Educ. XXXX, XXX, XXX−XXX

general chemistry by U.S. high school courses because, in part, our mean pretest score of 44.4% on the Chemical Concepts Inventory is similar to Mulford and Robinson's original 45.5%.[43] Given that alternate conceptions are common among students entering college general chemistry, it would be ideal to prevent them from forming in the first place. Perhaps the best way to prevent the formation of alternate conceptions is to elicit a class set of student conceptions—in agreement with the modern scientific consensus or not—and treat all of them as potential hypotheses to be tested against data. This mirrors the process of science itself. Additionally, it allows the introduction of cognitive conflict and resolution through social construction, both structural elements of a learning cycle curriculum, which is discussed in more detail in the following.

With respect to scientific reasoning ability, the results of this study are corroborated by the results of the study published in 2007 by Lewis and Lewis.[23] They found that a pretest of scientific reasoning ability predicted which students were at-risk of not successfully completing the general chemistry sequence. These combined results imply that development of students' scientific reasoning abilities should be a curricular goal for college preparatory chemistry courses. There have been successes in curriculum design that have led to the improvement of students' scientific reasoning ability. Lawson et al.[59] found that a learning cycle-based curriculum[8,60] in a college nonmajors biology course changed the scientific reasoning ability of students in the course from 9% formal (defined as >85% correct on a test of scientific reasoning) before the semester to a remarkable 44% formal by the end of the term.

Carmel, Jessa, and Yezierski[61] also found evidence of improvement of scientific reasoning skills in some students due to the design of a college chemistry curriculum in a course for nonmajors. As with Lawson et al.,[59] a learning-cycle-based curriculum was employed. In their study, 16 participants had potential to make gains in proportional reasoning ability, and nine demonstrated gains.

We believe that the key feature of curricula such as those of Lawson et al.[59] and Carmel, Jessa, and Yezierski,[61] which have been shown to promote the development of scientific reasoning skills, is a data-to-concepts curriculum sequence, often referred to more broadly as a learning cycle.[8,60] Not only is subject matter knowledge constructed in the mind of the learner,[9,10] but also procedural knowledge about scientific reasoning is constructed by the individual. When instructors present prepacked, predigested conceptual systems to students in lecture and textbooks, they take away the opportunity for students to construct knowledge. When the classroom and laboratory learning environments do not require students to think, their thinking skills cannot develop.

We must also be cognizant of the evidence that a preparatory course with a focus on only basic vocabulary, concepts, and problem-solving skills may detract from the national goal of improving the STEM pipeline. In a six-year study of the effect of a preparatory course on success in the first-semester general chemistry course, Bentley and Gellene[62] found that 44% of the students who passed a preparatory course with a C or better did not enroll in a general chemistry course in the subsequent semester, effectively reducing the total number of students who complete the first-semester general chemistry course and narrowing the STEM pipeline rather than expanding it.

If the problem of underdeveloped scientific reasoning skills continues to be not addressed, students will continue to be lost from the STEM pipeline. High school chemistry and college preparatory courses must be designed to promote the development of scientific reasoning skills within the context of the discipline. Both curriculum development and professional development efforts need to be undertaken so that chemistry curriculum and teaching become focused on both declarative knowledge *and* scientific reasoning skills construction.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available on the ACS Publications website at DOI: 10.1021/acs.jchemed.5b00146.

Histograms and frequency tables for scores from each instrument (PDF, DOCX)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: mark.cracolice@umontana.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Daniel Denis of the University of Montana provided much-appreciated expertise on the statistical analysis. We value our collaboration with Earle Adams, also of the University of Montana, who allowed us to codesign the general chemistry course laboratory schedule to accommodate time for cognitive testing. We are also grateful to our students for their cooperation. Additionally, the input from three anonymous reviewers was very valuable in shaping the final version of the article.

■ REFERENCES

(1) Tai, R. H.; Sadler, P. M.; Loehr, J. F. Factors Influencing Success in Introductory College Chemistry. *J. Res. Sci. Teach.* **2005**, *42*, 987–1012.

(2) Gillespie, R. J. What is Wrong with the General Chemistry Course? *J. Chem. Educ.* **1991**, *68*, 192–194.

(3) Freeman, S.; Eddy, S. L.; McDonough, M.; Smith, M. K.; Okoroafor, N.; Jordt, H.; Wenderoth, M. P. Active Learning Increases Student Performance in Science, Engineering, and Mathematics. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 8410–8415.

(4) Chambers, K. A.; Blake, B. The Effect of LearnStar on Student Performance in Introductory Chemistry. *J. Chem. Educ.* **2008**, *85*, 1395–1399.

(5) Gafney, L. Workshop Evaluation. In *Peer-Led Team Learning: A Guidebook*; Gosser, D. K., Cracolice, M. S., Kampmeier, J. A., Roth, V., Strozak, V. S., Varma-Nelson, P., Eds.; Prentice Hall: Upper Saddle River, NJ, 2001; pp 75–93.

(6) Nakhleh, M. B. Why Some Students Don't Learn Chemistry: Chemical Misconceptions. *J. Chem. Educ.* **1992**, *69*, 191–196.

(7) Herron, J. D. *The Chemistry Classroom: Formulas for Successful Teaching*; American Chemical Society: Washington, DC, 1996.

(8) Lawson, A. E.; Abraham, M. R.; Renner, J. W. *A Theory of Instruction: Using the Learning Cycle To Teach Science Concepts and Thinking Skills*; National Association for Research in Science Teaching: Cincinnati, OH, 1989.

(9) Bodner, G. Constructivism: A Theory of Knowledge. *J. Chem. Educ.* **1986**, *63*, 873–878.

(10) Cracolice, M. S. How Students Learn: Knowledge Construction in College Chemistry Courses. In *Chemists' Guide to Effective Teaching*; Pienta, N. J., Cooper, M. M., Greenbowe, T. J., Eds.; Prentice Hall: Upper Saddle River, NJ, 2005; pp 12–27.

DOI: 10.1021/acs.jchemed.5b00146
J. Chem. Educ. XXXX, XXX, XXX–XXX

(11) Niaz, M. *Teaching General Chemistry: A History and Philosophy of Science Approach*; Nova Science Publishers, Inc.: New York, 2008.
(12) Cornog, J.; Stoddard, G. D. Predicting Performance in Chemistry. *J. Chem. Educ.* **1925**, *2*, 701−708.
(13) ACS Division of Chemical Education Examinations Institute. *Exams Page*; American Chemical Society: Washington, DC, 2015. http://chemexams.chem.iastate.edu/instructors/assessment-materials/exams (accessed May 21, 2015).
(14) Russell, A. A. A Rationally Designed General Chemistry Diagnostic Test. *J. Chem. Educ.* **1994**, *71*, 314−317.
(15) Hovey, N. W.; Krohn, A. Predicting Failures in General Chemistry. *J. Chem. Educ.* **1958**, *35*, 507−509.
(16) Hovey, N. W.; Krohn, A. An Evaluation of the Toledo Chemistry Placement Examination. *J. Chem. Educ.* **1963**, *40*, 370−372.
(17) California State Board of Education. *Science Content Standards for California Public Schools: Kindergarten through Grade Twelve*; California Department of Education: Sacramento, CA, 1988.
(18) Neisser, U.; Boodoo, G.; Bouchard, T. J., Jr.; Boykin, A. W.; Brody, N.; Ceci, S. J.; Halpern, D. F.; Loehlin, J. C.; Perloff, R.; Sternberg, R. J.; Urbina, S. Intelligence: Knowns and Unknowns. *Am. Psychol.* **1996**, *51*, 77−101.
(19) Garnett, P. J.; Garnett, P. J.; Hackling, M. W. Students' Alternative Conceptions in Chemistry: A Review of Research and Implications for Teaching and Learning. *Stud. Sci. Educ.* **1995**, *25*, 69−95.
(20) Pfundt, H.; Duit, R. *Bibliography: Students' Alternative Frameworks and Science Education*; University of Kiel Institute for Science Education: Kiel, Germany, 2000.
(21) Wandersee, J. H.; Mintzes, J. J.; Novak, J. D. Research on Alternative Conceptions in Science. In *Handbook of Research on Science Teaching and Learning*; Gabel, D. L., Ed.; Macmillan: New York, 1994; pp 177−210.
(22) Lawson, A. E. What Do Tests of "Formal" Reasoning Actually Measure? *J. Res. Sci. Teach.* **1992**, *29*, 965−983.
(23) Lewis, S. E.; Lewis, J. E. Predicting At-Risk Students in General Chemistry: Comparing Formal Thought to a General Achievement Measure. *Chem. Educ. Res. Pract.* **2007**, *8*, 32−51.
(24) Bird, L. Logical Reasoning Ability and Student Performance in General Chemistry. *J. Chem. Educ.* **2010**, *87*, 541−545.
(25) BouJaoude, S.; Salloum, S.; Abd-El-Khalick, F. Relationships Between Selective Cognitive Variables and Students' Ability to Solve Chemistry Problems. *Int. J. Sci. Educ.* **2004**, *26*, 63−84.
(26) Cracolice, M. S.; Deming, J. C.; Ehlert, B. Concept Learning Versus Problem Solving: A Cognitive Difference. *J. Chem. Educ.* **2008**, *85*, 873−878.
(27) Modestou, M.; Gagatsis, A. Cognitive and Metacognitive Aspects of Proportional Reasoning. *Math. Think. Learn.* **2010**, *12*, 36−53.
(28) Chan, J. Y. K.; Bauer, C. F. Identifying At-Risk Students in General Chemistry via Cluster Analysis of Affective Characteristics. *J. Chem. Educ.* **2014**, *91*, 1417−1425.
(29) Lewis, S. E.; Shaw, J. L.; Heitz, J. O.; Webster, G. H. Attitude Counts: Self-Concept and Success in General Chemistry. *J. Chem. Educ.* **2009**, *86*, 744−749.
(30) Inhelder, B.; Piaget, J. *The Growth of Logical Thinking*; Basic Books: New York, 1958.
(31) Kwon, Y.-J.; Lawson, A. E. Linking Brain Growth with the Development of Scientific Reasoning Ability and Conceptual Change During Adolescence. *J. Res. Sci. Teach.* **2000**, *37*, 44−62.
(32) Lawson, A. E. A Review of Research on Formal Reasoning and Science Teaching. *J. Res. Sci. Teach.* **1985**, *22*, 569−617.
(33) Bao, L.; Cai, T.; Koenig, K.; Fang, K.; Han, J.; Wang, J.; Liu, Q.; Ding, L.; Cui, L.; Luo, Y.; Wang, Y.; Li, L.; Wu, N. Learning and Scientific Reasoning. *Science* **2009**, *323*, 586−587.
(34) 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*, 535−549.
(35) Shayer, M.; Ginsburg, D.; Coe, R. Thirty Years On—A Large Anti-Flynn Effect? The Piagetian Test *Volume & Heaviness* Norms 1975−2003. *Br. J. Educ. Psychol.* **2007**, *77*, 25−41.
(36) Abraham, M. R.; Pavelich, M. R. *Inquiries into Chemistry*, 3rd ed.; Waveland Press: Long Grove, IL, 1999.
(37) Lewis, S. E.; Lewis, J. E. Departing from Lectures: An Evaluation of a Peer-Led Guided Inquiry Alternative. *J. Chem. Educ.* **2005**, *82*, 135−139.
(38) Gosser, D. K.; Cracolice, M. S.; Kampmeier, J. A.; Roth, V.; Strozak, V. S.; Varma-Nelson, P. *Peer-Led Team Learning: A Guidebook*; Prentice Hall: Upper Saddle River, NJ, 2001.
(39) Lewis, S. E. Retention and Reform: An Evaluation of Peer-Led Team Learning. *J. Chem. Educ.* **2011**, *88*, 703−707.
(40) Bergmann, J.; Sams, A. *Flip Your Classroom: Reach Every Student in Every Class Every Day*; International Society for Technology in Education: Eugene, OR, 2012.
(41) Strayer, J. F. How Learning in an Inverted Classroom Influences Cooperation, Innovation, and Task Orientation. *Learning Environ. Res.* **2012**, *15*, 171−193.
(42) Gall, M. D.; Borg, W. R.; Gall, J. P. *Educational Research: An Introduction*, 6th ed.; Longman Publishers USA: White Plains, NY, 1996.
(43) Mulford, D. R.; Robinson, W. R. An Inventory for Alternate Conceptions Among First-Semester General Chemistry Students. *J. Chem. Educ.* **2002**, *79*, 739−744.
(44) Barbera, J. A Psychometric Analysis of the Chemical Concepts Inventory. *J. Chem. Educ.* **2013**, *90*, 546−553.
(45) Bramaje, G. P.; Espinosa, A. A. Peer-Led Team Learning Approach: Effects on Students' Conceptual Understanding and Attitude Towards Chemistry. *Int. J. Learn. Teach.* **2013**, *5*, 55−77.
(46) Kruse, R. A.; Roehrig, G. H. A Comparison Study: Assessing Teachers' Conceptions with the Chemistry Concepts Inventory. *J. Chem. Educ.* **2005**, *82*, 1246−1250.
(47) Raven, J.; Raven, J. C.; Court, J. H. *Manual for Raven's Progressive Matrices and Vocabulary Scales*; Harcourt Assessment: San Antonio, TX, 1998.
(48) *Raven's Standard Progressive Matrices (SPM): Evidence of Reliability and Validity*; NCS Pearson, Inc.: Bloomington, MN, 2007.
(49) Lawson, A. E. The Development and Validation of a Classroom Test of Formal Reasoning. *J. Res. Sci. Teach.* **1978**, *15*, 11−24.
(50) Jensen, J. L.; McDaniel, M. A.; Woodard, S. M.; Kummer, T. A. Teaching to the Test···or Testing to Teach: Exams Requiring Higher Order Thinking Skills Encourage Greater Conceptual Understanding. *Educ. Psychol. Rev.* **2014**, *26*, 307−329 (2014).
(51) Lawson, A. E.; Alkhoury, S.; Benford, R.; Clark, B. R.; Falconer, K. A. What Kinds of Scientific Concepts Exist? Concept Construction and Intellectual Development in College Biology. *J. Res. Sci. Teach.* **2000**, *37*, 996−1018.
(52) Chletsos, P. N.; De Lisi, R.; Turner, G.; McGillicuddy-De Lisi, A. V. Cognitive Assessment of Proportional Reasoning Strategies. *J. Res. Dev. Educ.* **1989**, *22*, 18−27.
(53) Bauer, C. F. Attitude Towards Chemistry: A Semantic Differential Instrument for Assessing Curriculum Impacts. *J. Chem. Educ.* **2008**, *85*, 1440−1445.
(54) Xu, X.; Lewis, J. E. Refinement of a Chemistry Attitude Measure for College Students. *J. Chem. Educ.* **2011**, *88*, 561−568.
(55) ACS Division of Chemical Education Examinations Institute. *First Term General Chemistry Form 2009*; ACS DivCHED Examinations Institute, Iowa State University: Ames, IA, 2009.
(56) ACS Division of Chemical Education Examinations Institute. http://chemexams.chem.iastate.edu/sites/default/files/national-norms/gc09f.html (accessed May 21, 2015).
(57) Cowles, M.; Davis, C. On the Origins of the 0.05 Level of Statistical Significance. *Am. Psychol.* **1982**, *37*, 553−558.
(58) Cohen, J. A Power Primer. *Psychol. Bull.* **1992**, *112*, 155−159.
(59) Lawson, A. E.; Clark, B.; Cramer-Meldrum, E.; Falconer, K. A.; Sequist, J. M.; Kwon, Y.-J. Development of Scientific Reasoning in College Biology: Do Two Levels of General Hypothesis-Testing Skills Exist? *J. Res. Sci. Teach.* **2000**, *37*, 81−101.

DOI: 10.1021/acs.jchemed.5b00146
J. Chem. Educ. XXXX, XXX, XXX−XXX

(60) Cracolice, M. S. Guided Inquiry and the Learning Cycle. In *Chemists' Guide to Effective Teaching*; Pienta, N. J., Cooper, M. M., Greenbowe, T. J., Eds.; Prentice Hall: Upper Saddle River, NJ, 2009; Vol. *II*, pp 20−34.
(61) Carmel, J. H.; Jessa, Y.; Yezierski, E. J. Targeting the Development of Content Knowledge and Scientific Reasoning: Reforming College-Level Chemistry for Nonscience Majors. *J. Chem. Educ.* **2015**, *92*, 46−51.
(62) Gellene, G. I.; Bentley, A. B. A Six-Year Study of the Effects of a Remedial Course in the Chemistry Curriculum. *J. Chem. Educ.* **2005**, *82*, 125−130.

DOI: 10.1021/acs.jchemed.5b00146
J. Chem. Educ. XXXX, XXX, XXX−XXX