JOURNAL OF CHEMICAL EDUCATION

Article

pubs.acs.org/jchemeduc

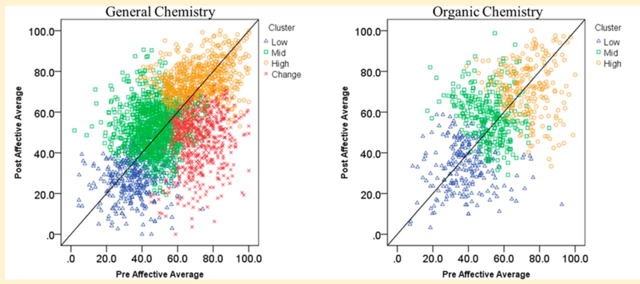
Measuring Meaningful Learning in the Undergraduate Chemistry Laboratory: A National, Cross-Sectional Study

Kelli R. Galloway and Stacey Lowery Bretz*

Department of Chemistry & Biochemistry, Miami University, Oxford, Ohio 45056, United States

Supporting Information

ABSTRACT: Research on laboratory learning points to the need to better understand what and how students learn in the undergraduate chemistry laboratory. The Meaningful Learning in the Laboratory Instrument (MLLI) was administered to general and organic chemistry students from 15 colleges and universities across the United States in order to measure the students' cognitive and affective expectations and experiences within the context of performing experiments in their chemistry laboratory courses. Data were analyzed using exploratory factor analysis and cluster analysis. The factor analysis revealed unique mental frameworks for how students think about their laboratory experiences. Exploration of the cluster analysis output indicated a four cluster solution for general chemistry students and a three cluster solution for organic chemistry students. The clusters were further analyzed by examining item pre versus post scatterplots to characterize their unique cognitive and affective expectations and experiences for learning. Both courses had a cluster of students with high cognitive and affective expectations that were fulfilled by their laboratory experiences, as well as a cluster of students who had high cognitive expectations but low affective expectations. This cluster's cognitive expectations went unfulfilled, while their negative affective expectations were fulfilled, and their disparate cognitive and affective perceptions created a hindrance for the necessary integration of cognitive, affective, and psychomotor domains for meaningful learning.



KEYWORDS: *Chemistry Education Research, Testing/Assessment, Laboratory Instruction, First-Year Undergraduate/General, Second-Year Undergraduate, Organic Chemistry, Hands-On Learning/Manipulatives, Inquiry-Based/Discovery Learning, Learning Theories, Problem Solving/Decision Making*

FEATURE: Chemical Education Research

INTRODUCTION

The undergraduate chemistry laboratory is an integral component of undergraduate chemistry education, but in the decades of published chemistry education research, relatively few studies offer insight into what and how students learn within their chemistry laboratory courses.[1−6] The undergraduate chemistry laboratory has long been argued as an important area for continued research and reform as numerous reviews have called for rigorous research to investigate student learning in the university laboratory.[3−10] The literature includes reports regarding the effects of a pre-lab lecture vs post-lab lecture,[11] peer-led learning,[12] incorporation of simulations,[13−15] and students' perceptions of the use of instruments in the laboratory.[16,17] Other research reports have investigated the effectiveness of new laboratory curricula including Model-Observe-Reflect-Explain (MORE),[18,19] problem-based learning experiments,[20] and cooperative learning[21−23] (to name a few). Additional research has examined the impacts of research-based undergraduate laboratory curricula including the CASPiE project[24−26] and other research-based curricula.[27] The body of research on the undergraduate chemistry laboratory includes reports regarding general chemistry courses,[11,12,15,18−23] organic chemistry,[13] and upper-division chemistry courses, including inorganic qualitative analysis[15] and instrumental analysis.[16,17]

These previous studies have identified some important aspects of student learning in the laboratory. Jalil found that when students conduct an experiment prior to a lecture or discussion of the concepts, the students self-report increased understanding of the content, enjoyment of the lab, and the ability to achieve in the lab as the students were given the opportunity to develop independence and self-reliance in the laboratory.[11] A study investigating the effectiveness of the MORE laboratory curriculum found that when students were explicitly asked to draw molecular diagrams of their mental frameworks and then reflect on those drawings after conducting an experiment, students revised their models to be consistent with scientifically accurate ideas.[18] However, a follow-up study reported that when the students in the MORE laboratory were given a slightly different context for the molecular drawing, they could not transfer their knowledge gained from reflection to the new context.[19] Thus, there must have been additional factors

DOI: 10.1021/acs.jchemed.5b00538
J. Chem. Educ. XXXX, XXX, XXX−XXX

that influenced how students constructed meaning from their laboratory experiments. The development and incorporation of problem-based, cooperative learning experiments at one university organized students into project teams.[21] Assessment of the curriculum found increased metacognitive awareness, as well as evidence for students taking charge and figuring out for themselves how to engage in their learning experience.[22,23]

To date, there are some notable gaps in the literature regarding learning in the undergraduate chemistry laboratory. While the obvious emphasis of laboratory instruction is to teach hands-on skills and manipulation of equipment with precision and accuracy, the psychomotor domain is but one dimension of learning.[28−30] How do students cognitively process during experiments? What is the role of the affective domain in laboratory learning? How do cognitive, affective, and psychomotor learning integrate to catalyze meaningful learning experiences for students?[29,30] There are also methodological limitations. While some of these studies involved multiyear data collections, the majority of the studies described above were confined to investigating the phenomenon of interest within a single course at a single university.

■ THEORETICAL FRAMEWORK

In 2012, the National Research Council published the report on Discipline-Based Education Research (DBER) that compiled and synthesized empirical research across the fields of science and engineering education research.[31] The report concluded with a research agenda to further the understanding of how people learn within the disciplines of science and engineering. Of the 10 research recommendations laid out in the report, 4 recommendations in particular shaped the design of our study to investigate students' perceptions of their learning in the undergraduate chemistry laboratory:

- Research to understand how students learn in the laboratory setting
- Instruments to assess skills that well-designed laboratory instruction can promote
- Rigorous research to examine outcomes associated with the affective domain
- Investigations across multiple courses including cross-sectional studies

To investigate how students learn in the laboratory, including the role of the affective domain, we framed our research questions and methodological choices using Novak's theory of meaningful learning and human constructivism as a guiding theoretical lens.[28−30] For meaningful learning to occur, the learner must have relevant prior knowledge, the new knowledge must be presented in a meaningful way to the learner, and the learner must actively choose to engage in the learning process.[28−30,32] Given the learner's responsibility to choose to relate new knowledge to prior knowledge, the learner plays an active role in constructing knowledge through experiences.[28,32] Novak classifies these human experiences as cognitive (thinking), affective (feeling), and psychomotor (doing).[28−30] The successful integration of the cognitive, affective, and psychomotor experiences then leads to meaningful learning.[28]

We have previously reported the development of the Meaningful Learning in the Laboratory Instrument (MLLI).[33] As students take part in the psychomotor "doing" of learning in the undergraduate chemistry laboratory, the MLLI measures students' expectations for and experiences with the cognitive and affective domains of learning.[33] MLLI consists of 30 items and 1 indicator item as a reading check. Each of the 30 items inquires about a cognitive and/or affective experience that the student could have during their laboratory course. Students are asked to indicate their percent agreement (0% as Complete Disagree to 100% as Completely Agree) with each statement. The MLLI is administered at the beginning of the semester prior to completing any laboratory work in order to capture students' expectations for learning. At the end of the semester, the MLLI is administered again (with the verb tense of the item changed from future to past tense) in order to ask students about their learning experiences in the laboratory. Students' responses are compared between the beginning and the end of semester to explore how their experiences aligned with their expectations. Items on the MLLI are coded using the meaningful learning framework as cognitive, affective, or cognitive/affective (containing both cognitive and affective parts in the item).

We have reported findings with the MLLI regarding general and organic chemistry students' cognitive and affective expectations/experiences at one university.[33] We were curious to know whether (and if so, how) students from other universities in different learning environments would respond to the MLLI. Therefore, a study was conducted using the MLLI to measure students' expectations and experiences within a cross-sectional data set of general and organic chemistry students at 15 additional colleges and universities. The goal of this study was to investigate how the MLLI functioned for a much larger sample of general and organic chemistry students, as well as to characterize the cognitive and affective expectations/experiences for laboratory learning for these students.

■ RESEARCH QUESTIONS

The research questions for this study were (1) what are the cognitive and affective expectations and experiences for general and organic chemistry students? (2) From this information, what evidence, if any, points toward the integration of the cognitive and affective domains to create meaningful learning?

■ METHODS

School Recruitment

As the goal of this study was to explore how MLLI functioned with a larger number of students in a variety of laboratory environments, a sample of colleges and universities was recruited from across the United States. While criterion sampling could have been used to account for specific variables such as size of the institution or the chemistry department, location of the university, or a particular laboratory pedagogy, instead, purposeful sampling was used to recruit laboratory instructors who self-identified as interested in knowing more about their students' cognitive and affective experiences in the laboratory. The goal of the study was not to make a generalizable statement about the current state of undergraduate chemistry laboratories, but rather to analyze how MLLI functioned at different colleges and universities with different students in different learning environments.

An e-mail invitation to chemistry faculty was issued through an international chemistry education listserv. The invitation described the project, the MLLI, and the courses of interest for data collection. Interested laboratory instructors were asked to respond with information about their course and institution. Additional announcements and invitations to participate were

DOI: 10.1021/acs.jchemed.5b00538
J. Chem. Educ. XXXX, XXX, XXX−XXX

also extended at national conferences. Ultimately, 26 instructors from different colleges and universities around the continental United States indicated interest in participating in this data collection. IRB approval agreements were obtained at 17 of the 26 schools prior to the start of the fall 2014 semester; data collection proceeded with these 17 schools. Two schools dropped out due to poor student participation. Table 1 shows

Table 1. Participating Colleges and Universities

Institution Type	Private	Public	Total
Community college	0	1	1
Primarily undergraduate institution	1	0	1
Comprehensive university	0	5	5
Research university	2	6	8
Total	**3**	**12**	**15**

the demographic information for the remaining 15 colleges and universities from 11 states where MLLI data was collected and analyzed. Data was collected in general chemistry (GC) from 11 schools and in organic chemistry (OC) from 7 schools (some schools collected both GC and OC data).

Data Collection

The MLLI was administered online using Qualtrics Survey software, and individualized survey links were prepared for each school. Prior to the beginning of the fall 2014 semester, course instructors were given details for how to administer the MLLI to their students, including language to accompany the survey link. Instructors were told to send the MLLI to their students during the first week of class and that it could be included as part of the lab check-in procedures. Students were given 2 weeks to complete the assessment and were encouraged to do so prior to their first laboratory experiment of the semester. As the MLLI pre-test measures students' expectations for learning in their laboratory course, it is important for students to complete the survey prior to their "doing" of laboratory work as their expectations would then be influenced by their initial experience.

After the midterm of the semester, the lab instructors were provided instructions to administer the MLLI post-test. Again, individualized links were created for each school. The instructors were asked to send the MLLI to their students in mid-November to permit the students to respond during the 2 weeks prior to the Thanksgiving holiday.

Students were asked to provide their university e-mail address in order to facilitate the matching of their responses from the pre- to the post-test. This matching of responses allowed for analysis of a direct comparison between students' expectations and their actual experiences. While over 9500 students responded to the MLLI at least once during the fall 2014 semester from a variety of first and second year general and organic chemistry courses, this manuscript focuses upon the analysis of students with matched responses enrolled in only the first-semester general chemistry (GC) (N = 2814) or first-semester organic chemistry laboratory (OC) courses (N = 769). A few schools offered a one semester first-year chemistry laboratory course or a one semester organic chemistry laboratory course ("survey course"). Students in these courses were grouped together with the students in the first semester of the corresponding two semester sequence because the survey course would still constitute the first time that students would experience the material. Table 2 shows the distribution of students by institution type and course. This sample is weighted slightly heavier toward females (51% of GC and 63% of OC) and is ≥80% white.

Table 2. Distribution of Students by Institution Type and Course

Institution Type	Course (N) GC	OC	Total
Community college	22	0	22
Primarily undergraduate institution	0	24	24
Comprehensive university	190	6	196
Research university	2602	739	3341
Total	**2814**	**769**	**3583**

Analysis

To answer the research questions, the analysis needed to characterize students' cognitive and affective expectations and experiences. To begin, descriptive statistics were calculated by course, and a variety of plots were constructed to visualize the data. These plots included histograms of the composite scores for each of the cognitive, affective, and cognitive/affective domains for both the pre- and the post-test administrations. Scatterplots and boxplots were constructed for composite variables as well as for individual items to compare the responses from pre to post. While boxplots show the overall change from pre to post, scatterplots display how individuals change. A more complete understanding of the data can be constructed by comparing these different data visualizations side-by-side.

Given that the data set consisted of students' expectations and experiences across a variety of diverse colleges and universities, the internal structure of the MLLI was re-examined using Exploratory Factor Analysis (EFA). This technique grouped items together based on their inter-item correlations; thus, items that students answered similarly were grouped together.[34] For CER and other DBER fields, EFA is one way to explore students' mental models and frameworks.[35,36] The ideas that are grouped together in the mind of the student are communicated via their responses and can be detected by EFA.[37] This EFA was performed with all 30 items on four separate instances: both pre- and post-tests for both GC and OC. Items were reverse coded as needed prior to the EFA. Principal axis factoring was chosen as the factor extraction method due to deviations from normality for each item.[34] For factor rotation, a Promax rotation was chosen because it is an oblique rotation that allows for the factors to be correlated.[34] Meaningful learning would necessitate integration of the psychomotor with the cognitive and affective, so correlated factors would make theoretical sense.[28]

In addition to the EFA, the theoretical factors were analyzed by calculating a Cronbach α and an α "if item deleted" for both GC and OC for both the pre and post-test administrations.[38] These calculations allowed for an examination of the extent to which the items were correlated within the meaningful learning theoretical factors. Typically, α "if item deleted" is considered if the α would increase, suggesting that such an item would appear to be less strongly correlated to the other items in the factor. The interpretation of such a statistic when analyzing student knowledge frameworks or perceptions of learning is that if the item were to be deleted, then perhaps that item is not

DOI: 10.1021/acs.jchemed.5b00538
J. Chem. Educ. XXXX, XXX, XXX−XXX

Table 3. Descriptive Statistics for MLLI pre and post-test administration GC and OC

		Scale Mean (SD)					
		Cognitive		Affective		Cognitive/Affective	
Course	Measure	Pre	Post	Pre	Post	Pre	Post
GC	Mean (SD)[a]	69.9 (11.0)	58.1 (14.5)	57.1 (18.1)	53.2 (18.3)	56.2 (15.6)	43.4 (16.5)
N = 2814	α	0.77	0.79	0.80	0.76	0.64	0.62
OC	Mean (SD)[a]	68.5 (11.4)	60.0 (13.0)	54.0 (18.0)	54.7 (19.4)	52.8 (16.3)	49.5 (16.9)
N = 769	α	0.78	0.78	0.80	0.80	0.67	0.67

[a]Scale responses range from 0% to 100%.

as strongly connected in the mind of the student to the other items in that factor.[37] Items that were identified through this technique were analyzed further by examining scatterplots and boxplots to better understand students' response patterns for those items.

Hierarchical cluster analysis was conducted to identify natural student groups within each course.[39] Separate cluster analyses were conducted for each course. Hierarchical cluster analysis was chosen as the clustering technique because there was no theoretical justification to propose an a priori number of groups of students based on the MLLI responses. While a previous study conducted with the MLLI indicated 4 clusters of students at one institution, there was no prior evidence to expect a similar cluster solution with students from varying institutions across the country.[40] Cluster solutions were explored by considering students' composite scores for cognitive, affective, and cognitive/affective for the pre- and post-test administrations as well as without the cognitive/affective composite scores. The clustering variables were examined for collinearity by considering the correlations and variance inflation factor (VIF).[41] Correlations between the clustering variables ranged from 0.3 to 0.5, and the VIF values were less than 3.0 suggesting that there was no collinearity between these variables. (Correlation tables can be found in the Supporting Information). The distance measure was chosen as squared Euclidean distance because the clustering variables are a continuous measure and squared Euclidean distance can be interpreted as physical distances between points in Euclidean space.[39] Ward's method was chosen as the linkage method because it combines clusters by minimizing the sum of squared deviations within a group.[42,43] To identify the optimal cluster solution, the dendogram was analyzed by finding the "best cut" where the clusters below that cut were different from each other by the least amount.[41,43,44] For the dendogram, this cut is the break of the first large distance between clusters. The variance ratio criterion (VRC) was calculated to help identify the best solution.[41,45,46]

After considering all the possible cluster solutions, and in light of the parameters discussed in the previous paragraph, the results presented in this manuscript present and discuss the cluster solution without the cognitive/affective average variables, that is, students were grouped based solely on their separate cognitive and affective perceptions of learning in the laboratory. As the cognitive/affective items incorporate both cognitive and affective experiences, the overlap of these domains yielded clusters with somewhat less cohesion and separation which led to decision to cluster solely on the cognitive and affective scales. The clusters were analyzed with both descriptive and inferential statistics. Scatterplots and boxplots were constructed for each cluster to help describe the characteristics of the cluster from the other clusters. Separate one-way repeated-measures ANOVAs were run to analyze how each cluster's expectations compared with their actual experiences. The appropriate assumptions for repeated-measures ANOVA were examined, including independent random sampling, multivariate normal distributions, and homogeneity of variance. Normality was assessed with the Shapiro-Wilk test and the homogeneity of variance was assessed with Levene's Test. Both showed some deviation, but the repeated measures ANOVA is not as sensitive to departures from normality or to some heterogeneity of variance.[47] Students self-selected to participate in the study, so conclusions should be drawn carefully.

While the goal of both the EFA and the cluster analysis was to ultimately identify students' cognitive and affective expectations and experiences, interpretation of the results in a meaningful manner was not a straightforward process. Analysis of the clusters began by attempting to impose a 2 × 2 matrix upon each cluster: did the students expect a particular item on the MLLI (yes or no) and did they experience each item (yes or no). Thus, the comparison between expectation and experience could have been categorized by the cells in such a 2 × 2 as "expectation fulfilled," "expectation unfulfilled," "unexpected experience," or "didn't expect, didn't happen." However, analysis of the data revealed that item classifications for each cluster were not so simple. Given the distribution of responses (even within a cluster) on some items, it was not always possible to label an item as expected or experienced (or not). Furthermore, experiences as measured by some items contributed to meaningful learning, while other items would indicate such experiences hindered meaningful learning.

To consider both the distribution of responses and the contribution/inhibition of the item to meaningful learning, scatterplots for each item post versus pre were created for each cluster. These plots were analyzed to look for an overall general tendency for the cluster. Patterns across each cluster's responses revealed the degree and magnitude of the clusters' expectations and experiences. Some experiences were clearly expected and some expectations were clearly met. Likewise, there were some items that did not suggest an overall expectation or experience. Items with this result were insightful in characterizing the clusters. The results from these analyses are explained below.

■ RESULTS AND DISCUSSION

Descriptive Statistics

Table 3 summarizes the descriptive statistics for GC and OC students. Means and standard deviations were calculated for each scale (cognitive, affective, and cognitive/affective) for both the pre- and post-test administrations, as well as Cronbach α as one measure of internal consistency within each scale. Scatterplots were constructed to examine the relationship

DOI: 10.1021/acs.jchemed.5b00538
J. Chem. Educ. XXXX, XXX, XXX−XXX

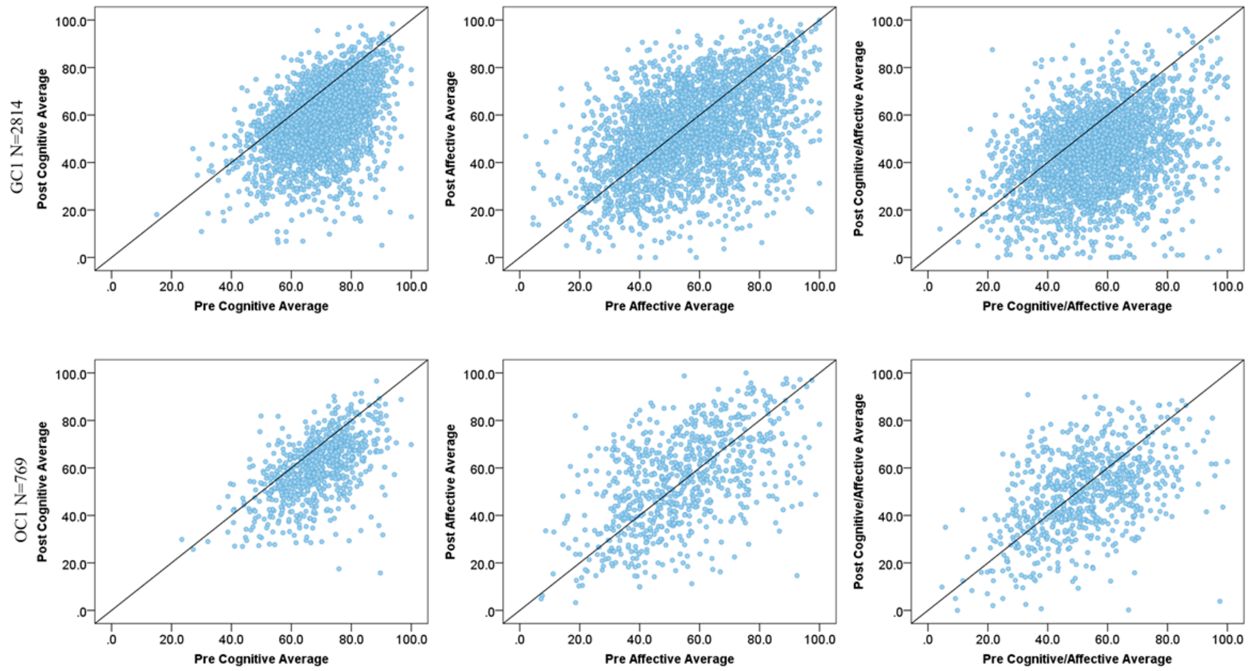


Figure 1. Scatterplots to compare GC and OC experiences versus expectations for cognitive, affective, and cognitive/affective scales.

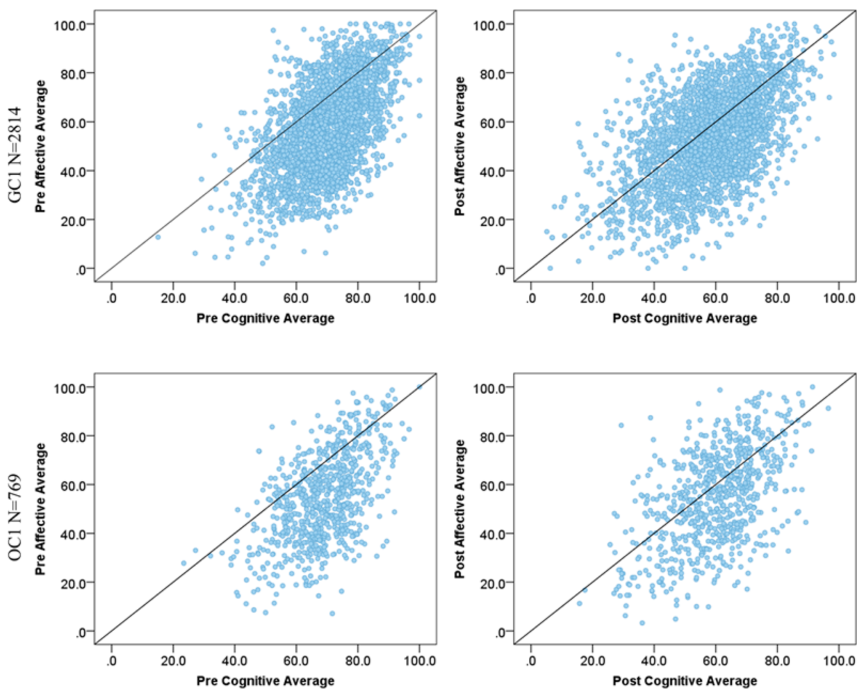


Figure 2. Scatterplots for GC (N = 2814) and OC (N = 769) comparing affective versus cognitive for pre and post separately.

between the pre and post responses for each scale (Figure 1). Each plot is drawn with a $y = x$ line for a visual examination of increase or decrease in responses from pre- to post-test.

There appear to be similar trends for both the GC and the OC students. For cognitive, there appears to be a narrow range of expectations with a wider range of reported cognitive experiences. The affective distributions are diverse for both expectations and experiences. There appear to be as many students whose affective experiences exceeded their expectations (points above the $y = x$ line) as there are students whose affective experiences did not meet their expectations (points below the $y = x$ line). The cognitive/affective plots have a greater spread than the cognitive plots but less than the affective plots. (Inferential statistics comparing GC and OC are presented in the Supporting Information.)

Scatterplots were also constructed to compare the cognitive and affective averages for pre and post separately (Figure 2). These plots also include a $y = x$ line which represents equal cognitive and affective responses. For both GC and OC, the plots for the pre-test show a narrow band of cognitive expectations mainly above 50% indicating high cognitive expectations, while the affective averages spanned the range

DOI: 10.1021/acs.jchemed.5b00538
J. Chem. Educ. XXXX, XXX, XXX−XXX

Table 4. Descriptive Statistics for Each Cluster for the Pre- and Post-Test Administrations

Cluster	Scale Mean (SD)					
	Cognitive		Affective		Cognitive/Affective	
	Pre	Post	Pre	Post	Pre	Post
Cluster 1 "Low" N = 401	60.3 (10.2)	43.3 (13.0)	36.6 (11.3)	28.3 (10.3)	42.8 (13.2)	26.8 (12.8)
Cluster 2 "Mid" N = 1176	67.7 (9.7)	60.4 (10.2)	48.2 (11.4)	52.7 (11.1)	50.3 (12.1)	43.2 (12.0)
Cluster 3 "High" N = 625	76.2 (9.2)	69.3 (11.6)	72.3 (13.2)	76.3 (8.4)	66.2 (13.3)	59.3 (13.4)
Cluster 4 "Change" N = 612	74.4 (9.8)	52.1 (13.5)	72.3 (10.1)	46.7 (13.2)	66.2 (12.9)	38.4 (14.7)

of the scale indicating diverse affective expectations for the sample. One interpretation of this data might be that despite their affective expectations for learning, the students expected to engage cognitively while conducting their laboratory experiments. For the post-test, both the cognitive and affective experiences are varied, with data spanning most of the possible range from 0 to 100.

Exploratory Factor Analysis

Fabrigar and Wegner discuss that when a common factor model is fitted to sample data, "it is extremely unlikely that the model will fit perfectly" (p 54).[34] Reasons they give for this include sample error resulting from the correlations calculated from the sample are unlikely identical to the population correlations as well as approximation error. Rather, a realistic goal is to strive for useful and appropriate statistical and conceptual utility. Instead of trying to identify an internal factor structure for MLLI based upon the national sample data, the goal of this analysis was to explore students' response patterns on the pre- and post-test administrations. The interpretability of the solutions was of greater importance than the commonly used statistical rules to determine the number of factors in MLLI. The scree plot was examined to identify the last large break between possible factors. For each course, the scree plot yielded two factors for both the pre and the post-tests. In addition to the two factor solution, the 3, 4, 5, and 6 factor solutions were also analyzed for their interpretability, and those results are presented in the Supporting Information.

The two factor solution was analyzed first as it was suggested by the scree plot. Interestingly, the items split into two groups based on how they affect meaningful learning. All the items worded such that a high score contributes to meaningful learning loaded onto one factor, while the items worded such that a high score hinders meaningful learning loaded onto the second factor. This phenomena has been documented in social psychology and could be attributed to a wording effect.[48,49] It is possible that the cognitive and affective domains are so disconnected in the students' minds that they only answer based on the positive and negative experiences they have in their chemistry laboratory courses. All but 3 of the 30 items loaded onto these two factors; the remaining three items (items 15, 17, and 26) had different behavior.

Q17 ("get stuck" but keep trying) and Q26 (make mistakes and try again) strongly loaded on both factors. They positively load on the positive scale but negatively load on the negative scale. These items are two-pronged, meaning they incorporate two ideas in the item. While "getting stuck" and making mistakes involve different aspects of a "wrong turn" in lab work (getting stuck means something is not working and not knowing what to do next, while making mistakes means carrying out a procedure incorrectly), the second part of both items involves a second chance. These items are not reverse coded for analysis because they involve an aspect of problem solving which contributes to meaningful learning.[28] One reason for loading on both factors could be that some students value working through these kinds of challenges in the laboratory, while other students become more easily frustrated and anxious when they get stuck or make mistakes. Another reason could be a result of the two-pronged nature of the items. Perhaps students are experiencing getting stuck and are making mistakes, but they are not trying again or working through their mistakes to find out how and why they got stuck in the first place.

The third item that did not load cleanly onto the positive or negative factors was Q15 (simple to do). For the pre-test EFA, Q15 did not load on either factor, while for the post-test EFA, Q15 loaded negatively on both factors. Similarly to Q17 and Q26, some students could view the simplicity of the experiments as a positive thing. They would not want to be challenged, and they prefer the ease of the experiments. On the other hand, some students may want to be challenged and feel bored in the laboratory because the experiments are *too* simple to do. This tension between the different perceptions of 'simple' likely contributed to the factor analysis results.

Cluster Analysis

To further explore GC and OC's cognitive and affective expectations and experiences for learning in the laboratory, cluster analysis was used to identify natural groupings of students within the sample. GC and OC students were clustered (separately) on their cognitive and affective pre- and post-test averages using a hierarchical agglomerative method with Squared Euclidean distance and Ward's linkage method. The cluster solution presented for each course is supported by the dendogram and agglomeration schedule, as well as by the interpretability and utility of the solution.[50,51] For both courses, a two cluster solution was best supported by the statistical output. However, meaningful interpretations of these clusters was limited at best. Therefore, the decision was made to explore additional solutions with more clusters in order to learn more about the student characteristics of the sample. Using the variance ratio criterion (VRC) developed by Calinski and Harabasz (1974), alternative solutions to the two cluster solutions were identified for both clusters.[41,45,46] The VRC was calculated for multiple solutions to identify a maximum. For GC, the next maximum after the two cluster solution was a four cluster solution. For OC, the VRC was maximized at the three cluster solution. These individual clusters were then explored for their unique characteristics to better understand the different perspectives of these students regarding their learning in the university chemistry laboratory. Additional visualizations for the cluster solutions are available in the Supporting Information.

GC Cluster Solution. Descriptive statistics for each cluster are presented in Table 4. To visualize the clusters, scatterplots and boxplots were constructed. In Figure 3, the initial

DOI: 10.1021/acs.jchemed.5b00538
J. Chem. Educ. XXXX, XXX, XXX−XXX

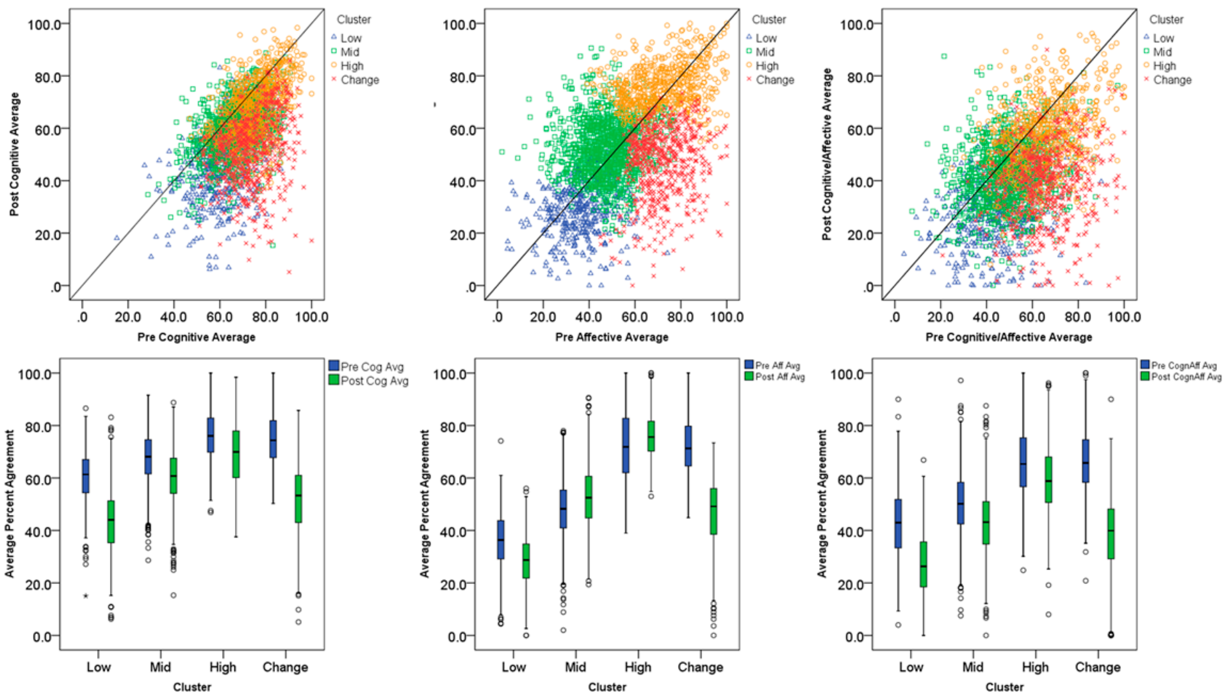


Figure 3. Scatterplots and boxplots for the four GC clusters.

Table 5. Results from Repeated Measures ANOVA Models for GC Clusters

Cluster	RM ANOVA Results Comparing Pre to Post on 3 Scales	Pairwise Comparisons (p, η^2)		
		Cognitive	Affective	Cogn/Aff
Low N = 401	Wilks's Λ = 0.358, $F(3, 398)$ = 237.6, $p \leq 0.0001$, η^2_p = 0.64	<0.0001, 0.60	<0.0001, 0.26	<0.0001, 0.48
Mid N = 1176	Wilks's Λ = 0.495, $F(3, 1173)$ = 399.2, $p \leq 0.0001$, η^2_p = 0.50	<0.0001, 0.32	<0.0001, 0.08	<0.0001, 0.18
High N = 625	Wilks's Λ = 0.517, $F(3, 622)$ = 193.7, $p \leq 0.0001$, η^2_p = 0.48	<0.0001, 0.28	<0.0001, 0.08	<0.0001, 0.16
Change N = 612	Wilks's Λ = 0.200, $F(3, 609)$ = 813.2, $p \leq 0.0001$, η^2_p = 0.80	<0.0001, 0.70	<0.0001, 0.74	<0.0001, 0.70

Table 6. Descriptive Statistics for OC Clusters for the Pre- and Post-Test Administrations

Cluster	Scale Mean (SD)					
	Cognitive		Affective		Cognitive/Affective	
	Pre	Post	Pre	Post	Pre	Post
Cluster 1 "Low" N = 232	61.5 (11.1)	51.1 (13.0)	38.1 (12.8)	33.3 (11.6)	39.9 (13.3)	34.1 (14.1)
Cluster 2 "Mid" N = 307	67.1 (8.9)	60.5 (10.7)	51.5 (10.8)	58.1 (11.8)	51.8 (11.6)	52.5 (12.1)
Cluster 3 "High" N = 230	77.4 (7.5)	68.4 (9.8)	73.2 (11.5)	71.7 (13.2)	66.2 (12.9)	61.2 (13.3)

scatterplots from Figure 1 were color-coded to indicate the clusters, and the boxplots show the overall tendency of each cluster. Of these four clusters, three clusters had sequential results: one cluster reported low expectations and low experiences, one cluster reported high expectations and high experiences, and one cluster fell in between these two. For these three clusters, the students' expectations appeared to shape their experiences, as shown by how each cluster remains in the same general area of the scale from their pre to post-test responses. The fourth cluster, however, reported higher expectations for laboratory yet experiences that did not fulfill those expectations. On both the scatterplots and boxplots, this fourth cluster had expectations similar to those of the high cluster, but reported experiences similar to those of the middle cluster. This exploration of the clusters led to assigning meaningful names to the clusters: Low (cluster 1), Mid (cluster 2), High (cluster 3), and Change (cluster 4).

To quantify the alignment between expectations and experiences as indicated by the students, one-way repeated measures ANOVA models were performed for each cluster (Table 5). The models indicated significant differences overall, and pairwise comparisons with a Bonferroni corrected p-value (α = 0.016) indicated significant differences for all three composite variables for each cluster. For cognitive, all four clusters showed large decreases from pre to post with Low and Change clusters having the largest effect sizes.[52,53] All clusters reported cognitive expectations for learning that went unfulfilled in their laboratory courses. For affective, Low and Change had large decreases from pre to post, but Mid and High showed moderate increases.[52,53] While Low and Change reported unfulfilled affective expectations, Mid and High's affective expectations were exceeded by their experiences. For cognitive/affective, all four clusters showed large decreases from pre to post, indicating unfulfilled expectations.[52,53]

DOI: 10.1021/acs.jchemed.5b00538
J. Chem. Educ. XXXX, XXX, XXX−XXX

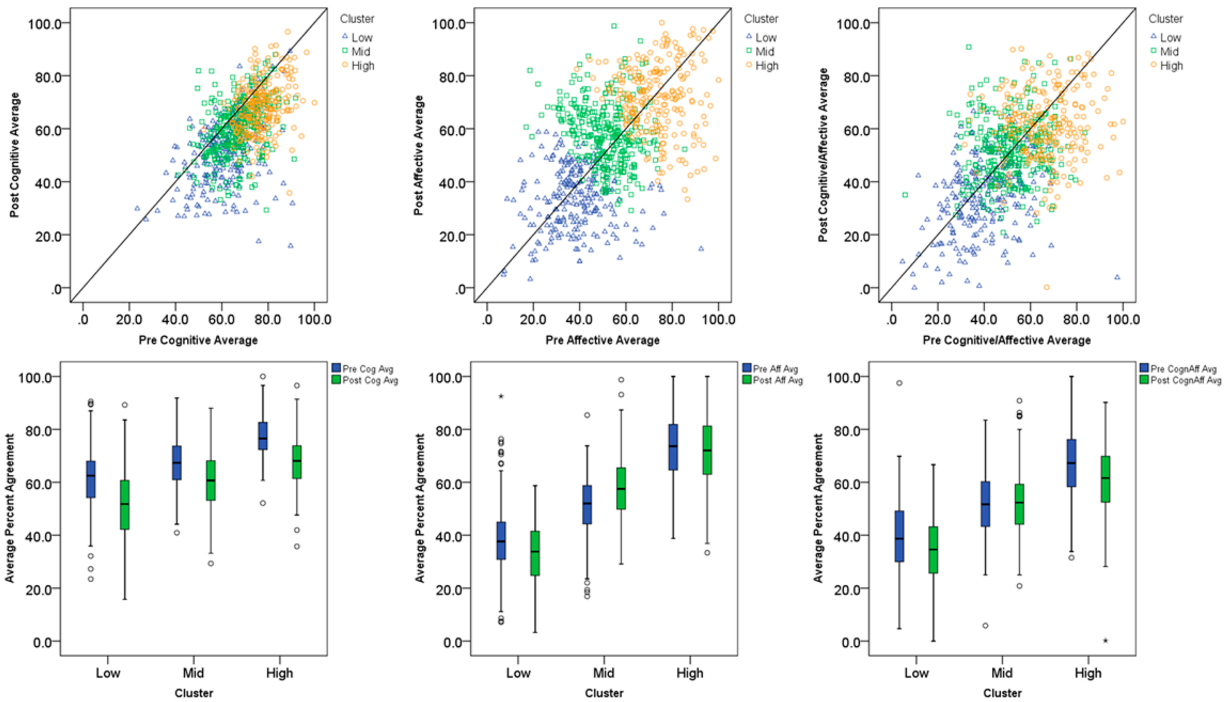


Figure 4. Scatterplots and boxplots for the three OC clusters.

Table 7. Results from Repeated Measures ANOVA for OC Clusters

		Pairwise Comparisons (p, η^2)		
Cluster	RM ANOVA Results Comparing Pre to Post on 3 Scales	Cognitive	Affective	Cogn/Aff
Low N = 232	Wilks's Λ = 0.607, $F(3, 229)$ = 49.6, $p \leq 0.0001$, η^2_p = 0.39	<0.0001, 0.38	<0.0001, 0.07	<0.0001, 0.11
Mid N = 307	Wilks's Λ = 0.600, $F(3, 304)$ = 67.6, $p \leq 0.0001$, η^2_p = 0.40	<0.0001, 0.27	<0.0001, 0.12	0.48
High N = 230	Wilks's Λ = 0.503, $F(3, 227)$ = 74.8, $p \leq 0.0001$, η^2_p = 0.50	<0.0001, 0.43	0.20	<0.0001, 0.09

Table 8. Expectations reported by cluster for GC and OC and their reported experience

Expectation	Did Experiences Fulfill This Expectation?					
MLLI item	GC Low	GC Mid	GC High	OC Low	OC Mid	OC High
Q3 to make decisions about what data to collect. (+C)	Mixed	Yes	Yes	Yes	Yes	Yes
Q5 to experience moments of insight. (+C)	No	Yes	Yes	Mixed	Yes	Yes
Q7 to learn critical thinking skills. (+C)	No	Yes	Yes	Mixed	Yes	Yes
Q10 to consider if my data makes sense. (+C)	Yes	Yes	Yes	Yes	Yes	Yes
Q13 to develop confidence in the laboratory. (+A)	No	Yes	Yes	Mixed	Yes	Yes
Q17 to "get stuck" but keep trying. (+C)	Yes	Yes	Mixed	Yes	Mixed	Mixed
Q19 to think about chemistry I already know. (+C)	Mixed	Yes	Yes	Mixed	Yes	Yes
Q22 to interpret my data beyond only doing calculations. (+C)	Mixed	Yes	Yes	Mixed	Yes	Yes
Q25 to use my observations to understand the behavior of atoms and molecules. (+C)	No	Mixed	Yes	Mixed	Mixed	Mixed
Q26 to make mistakes and try again. (+C)	Yes	Yes	Yes	Yes	Yes	Yes
Q31 to learn problem solving skills. (+C)	No	Mixed	Yes	Mixed	Yes	Yes

OC Cluster Solution. The descriptive statistics for the three cluster solution for OC students are shown in Table 6. These three clusters are sequential in nature, similar to three of the GC clusters. Analysis of their scatterplots and boxplots (Figure 4) revealed similar response patterns to Low, Mid, and High with the GC sample, so the OC clusters were assigned the same names.

One-way ANOVAs were conducted to examine the change from pre to post for each cluster (Table 7). Low had significant decreases from pre to post on all three variables, with medium and large effect sizes.[56,57] Mid had significant changes for cognitive and affective, but the cognitive averages decreased while the affective averages increased. Mid had no significant change for cognitive/affective. High had significant decreases for cognitive and cognitive/affective, but no significant change for affective.

Cluster Expectations and Experiences

Expectations Held by All Clusters. All clusters for both courses indicated they held high expectations for 11 of the 30 MLLI item laboratory experiences (Table 8). These items are all positively worded such that a high score contributes to meaningful learning. Ten of these 11 items are cognitive items, demonstrating the perspective students have to engage in cognitive processing in the laboratory. The remaining item

DOI: 10.1021/acs.jchemed.5b00538
J. Chem. Educ. XXXX, XXX, XXX−XXX

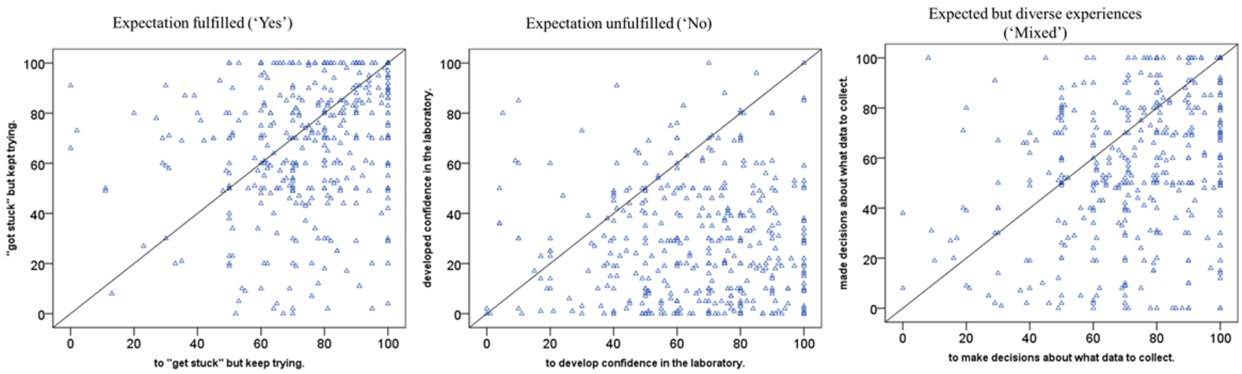


Figure 5. Scatterplots of items 17 ("get stuck" but keep trying), 13 (develop confidence in the laboratory), and 3 (make decisions about what data to collect) for GC Low ($N = 401$) to illustrate fulfilled expectations ("Yes"), unfulfilled expectations ("No"), and expectations with diverse experiences ('Mixed'). Since these plots are two-dimensional, item descriptives were used for a more accurate interpretation overall cluster behavior.

(Q13) is an affective item, indicating an expectation to grow in their confidence to do chemistry. Expectations that were clearly fulfilled for the majority of the cluster were marked "Yes" and expectations that were clearly unfulfilled for the majority of the cluster were marked "No". Expectations for a cluster who reported disparate experiences were marked "Mixed" (Figure 5).

While overall the High, Mid, and Low clusters tended to encounter the experiences they expected, there were interesting instances of unfulfilled expectations for each cluster at the item level. Looking at Table 8, no cluster reported meeting all 11 expectations. Neither High cluster indicated meeting their expectation to "get stuck" but keep trying (Q17). The GC High cluster reported meeting their expectation to use their observations to understand the behavior of atoms and molecules (Q25), but OC High did not. While both Mid clusters reported similar responses to the High clusters for these 11 expectations, both the Low clusters indicated that their experiences fulfilled fewer than half of these expectations. The only cognitive expectations that both Low clusters fulfilled were to "get stuck" but keep trying (Q17) and to make mistakes and try again (Q26). Additionally, the OC Low cluster indicated meeting their expectation to make decisions about what data to collect while this was not the case for the GC Low cluster. Both Low clusters' experiences failed to meet their expectations to experience moments of insight (Q5), learn critical thinking or problem solving skills (Q7 and Q31), develop confidence in the laboratory (Q13), think about chemistry they already know (Q19), interpret data beyond only calculations (Q22), or use observations to understand the behavior of atoms and molecules (Q25). Students from all clusters expected to engage in many cognitive experiences in the laboratory, but the students in the Low clusters were unable to fulfill them.

Different Expectations Across Clusters. In addition to the expectations expressed by all clusters in Table 8, each cluster reported expectations which distinguished them from the other clusters. Those different expectations, along with the reported cluster experience, are described below.

The High clusters had five additional expectations that were all positive in their contribution to meaningful learning (two affective items, two cognitive/affective, and one cognitive; Table 9). Both High clusters expected all of the positively worded MLLI experiences and did not expect any of the negatively worded MLLI experiences. The GC High cluster reported being excited to do chemistry (Q8) and confident using equipment (Q30). All but one of these 5 expectations were fulfilled for OC High. Neither High cluster reported experiencing that they thought about what the molecules were doing (Q11).

Table 9. Distinctive Expectations for GC and OC High Clusters

Expectation	Met by GC High?	Met by OC High?
Q1 to learn chemistry that will be useful in my life. (+C/A)	Mixed	Yes
Q8 to be excited to do chemistry. (+A)	Yes	Yes
Q11 to think about what the molecules are doing. (+C)	Mixed	Mixed
Q27 to be intrigued by the instruments. (+C/A)	Mixed	Yes
Q30 to be confident when using equipment. (+A)	Yes	Yes

Both Mid clusters indicated eight additional expectations beyond those in Table 8. Of these 8 expectations (Table 10), the 5 positive expectations were the same expectations as High's, and the 3 negative expectations were the same as Low's. As Mid falls between High and Low, the similarity of Mid's expectations to High and Low makes sense.

Table 10. Distinctive Expectations for GC and OC Mid Clusters

Expected by Mid	Met by GC Mid?	Met by OC Mid?
Q1 to learn chemistry that will be useful in my life. (+C/A)	Mixed	No
Q2 to worry about finishing on time. (−A)	Mixed	Mixed
Q8 to be excited to do chemistry. (+A)	Mixed	Mixed
Q11 to think about what the molecules are doing. (+C)	Mixed	No
Q14 to worry about getting good data. (−C/A)	Yes	Yes
Q20 to worry about the quality of my data. (−C/A)	Yes	Yes
Q27 to be intrigued by the instruments. (+C/A)	Mixed	Yes
Q30 to be confident when using equipment. (+A)	Yes	Yes

The Mid clusters only met half of their additional expectations. Both GC and OC clusters indicated meeting their expectations to worry about getting good data and the quality of their data (Q14 and Q20) and being confident using equipment (Q30). OC Mid, along with OC High, reported

DOI: 10.1021/acs.jchemed.5b00538
J. Chem. Educ. XXXX, XXX, XXX−XXX

meeting their expectation to be intrigued by the instruments. Perhaps the equipment used in the organic chemistry laboratory prompts increased curiosity and fascination that the equipment in the first year laboratory did not.

Both Low clusters reported seven additional expectations beyond Table 8. These seven expectations were all for negative experiences across a mixture of cognitive, affective, and cognitive/affective items (Table 11). Even though Low indicated expecting to engage cognitively during their laboratory experiences, they did not expect to have a positive affective experience. This misalignment between cognitive and affective expectations minimized opportunities for Low to engage in meaningful learning. Unfortunately, the Low clusters reported experiencing all seven negative expectations. Perhaps the act of expecting worry, nervousness, frustrations, intimidation, and confusion set in motion those feelings.[28]

Table 11. Distinctive Expectations for GC and OC Low Clusters

Expectations Met by GC and OC Low Clusters
Q2 to worry about finishing on time. (−A)
Q9 to be nervous about making mistakes. (−A)
Q14 to worry about getting good data. (−C/A)
Q20 to worry about the quality of my data. (−C/A)
Q21 to be frustrated. (−A)
Q28 to feel intimidated. (−A)
Q29 to be confused about what my data mean. (−C)

Interestingly, for three of Low's negative fulfilled expectations, the High clusters neither expected nor experienced them. Thus, the Low and High clusters reported nearly opposite perceptions of their laboratory learning. This finding not only speaks to the validity of the cluster solutions, but it also indicates the diversity of students present in the laboratory courses.

Disparate Expectations and Experiences. Analysis of the item scatterplots revealed that the responses to some items were quite varied rather than concentrated toward one corner or one side of the plot. The items with disparate responses show that while the cluster analysis identifies students with similarities, each student in a particular cluster will not necessarily respond to every item similarly. While many characteristics of the Low, Mid, and High clusters were similar across GC and OC, there were also differences in the indiscriminant items for both the GC and OC clusters.

For example, the responses of students in the High cluster in both GC and OC were quite varied for both Q14 (to worry about getting good data) and Q20 (to worry about the quality of my data). As discussed earlier, these students fulfilled the majority of their cognitive expectations for learning. They reported making decisions, having moments of insight, learning problem solving and critical thinking skills, and thinking about their prior chemistry knowledge. And yet, many students in the High clusters worried about obtaining quality data. This is further supported by examining the High clusters' responses to items in the EFA factor focused on the outcome of the experiment (Q2, Q9, Q10, Q14, and Q20) as their responses to all these questions were also varied (with the exception of item 10 as all clusters had a consensus response on Q10). The fact that a cluster of students reported overall positive learning experiences in the laboratory does not guarantee that they do not also have negative affective experiences which could hinder their overall meaningful learning experience.

GC Change Cluster. The GC cluster solution yielded a unique "Change" cluster whose expectations for their laboratory learning were not realized in their experiences (Table 12). As noted above, this cluster reported expectations similar to those of the High clusters, but reported experiences similar to those of students in the Mid and Low clusters. Interestingly, this cluster did not persist in the OC sample.

Change's fulfilled expectations were similar to those of High and Mid. All 5 fulfilled expectations were positive, as were their unmet expectations. This offers further evidence that Change had their sights set on enjoying their chemistry laboratory experience and engaging in meaningful learning activities. However, all of Change's unexpected experiences were negative, suggesting that Change was unprepared to navigate the gap between what they wanted for their learning and what they experienced. There are multiple possible reasons why such a gap existed. Perhaps the students in the Change cluster had little to no prior knowledge upon which to form expectations for learning in the laboratory. That lack of prior knowledge could have prevented these students from assimilating new experiences. Or perhaps the prior knowledge of these students regarding laboratory learning was not aligned with the laboratory curriculum. Additional research is underway to

Table 12. Expectations versus Experiences for GC Change Cluster

Fulfilled Expectations	Unfulfilled Expectations	Unexpected Experiences
Q3 to make decisions about what data to collect. (+C)	Q1 to learn chemistry that will be useful in my life. (+C/A)	Q2 to worry about finishing on time. (−A)
Q10 to consider if my data makes sense. (+C)	Q5 to experience moments of insight. (+C)	Q4 to feel unsure about the purpose of the procedures. (−C/A)
Q17 to "get stuck" but keep trying. (+C)	Q7 to learn critical thinking skills. (+C)	Q6 to be confused about how the instruments work. (−C)
Q26 to make mistakes and try again. (+C)	Q8 to be excited to do chemistry. (+A)	Q9 to be nervous about making mistakes. (−A)
Q30 to be confident when using equipment. (+A)	Q11 to think about what the molecules are doing. (+C)	Q12 to feel disorganized. (-C/A)
	Q13 to develop confidence in the laboratory. (+A)	Q21 to be frustrated. (−A)
	Q19 to think about chemistry I already know. (+C)	Q24 to focus on procedures, not concepts. (−C)
	Q22 to interpret my data beyond only doing calculations. (+C)	Q28 to feel intimidated. (−A)
	Q25 to use my observations to understand the behavior of atoms and molecules. (+C)	Q29 to be confused about what my data mean. (−C)
	Q27 to be intrigued by the instruments. (+C/A)	
	Q31 to learn problem solving skills. (+C)	

DOI: 10.1021/acs.jchemed.5b00538
J. Chem. Educ. XXXX, XXX, XXX−XXX

investigate whether the specific laboratory curricula for the students in this study had any effects on their MLLI responses.

CONCLUSIONS

The MLLI measured a range of cognitive and affective expectations and experiences for laboratory learning among GC and OC students (N = 3583) at 15 institutions across the United States. A nontraditional use of exploratory factor analysis suggested lenses through which students thought about their learning in the chemistry laboratory. Cluster analysis showed distinct natural groups of students within both the GC and OC samples. Both courses contained a cluster responding with low overall expectations and experiences, one cluster with high overall expectations and experiences, and a third cluster falling between the High and the Low clusters. These three clusters had overall expectations for learning that shaped the students' experiences.[54,55] The GC sample contained a fourth cluster, Change, whose expectations for laboratory learning were misaligned with their experiences, despite their high expectations. Previous cluster analysis using MLLI results from a local sample also yielded a four cluster solution with similar characteristics.[40] Because this data collection involved a much larger sample of students from diverse learning environments, similar cluster analysis results were not expected, and yet interesting to find. A comparison between the GC and OC responses showed MLLI could measure different expectations and experiences across groups of students whose prior knowledge of learning chemistry in the laboratory differs. While the Low, Mid, and High clusters for both GC and OC had similar expectations for learning, they did not report identical experiences.

This study responds to the research agenda put forth in the NRC DBER Report[31] by

- Furthering research on student learning in the laboratory setting by exploring student perceptions of their learning using a learning theory framework as a lens for research design and analysis
- Examining the integration of cognitive and affective domains of learning with the psychomotor domain
- Using a novel assessment tool designed to measure meaningful learning within the undergraduate chemistry laboratory
- Conducting the investigation across GC and OC through a cross-sectional research design at the national level

While additional research is still needed to understand how students learn in the laboratory setting in order to design evidence-based laboratory curriculum, this study puts forth evidence that the need to improve laboratory learning across the country is not unique to one school. Analysis is currently underway to explore how students responses on the MLLI varied in response to both faculty goals for learning in the laboratory and different laboratory curriculum, but it is beyond the scope of this study to draw conclusions regarding the role of pedagogy at any one school leading to increased (or decreased) learning experiences.

Implications for Teaching

The MLLI can be a useful tool for laboratory instructors at a variety of undergraduate institutions to measure students' perceptions of their learning. Results from the MLLI can reveal the expectations that students bring to the laboratory course and whether those expectations were realized in their laboratory experiences. In turn, the results can be used to inform the modification of laboratory curriculum and/or pedagogy for the following year. Faculty can examine the extent to which students meet their expectations for not only cognitive learning, but also affective learning.

High cognitive expectations for learning are necessary, but not sufficient, for meaningful learning. For the Low clusters, overall high expectations for cognitive learning experiences were accompanied by negative affective expectations. The affective profiles for all the clusters point to the need to better incorporate the affective domain into the design of laboratory curriculum. The incorporation of the affective domain ought to go beyond only attention to group work and connections to the real world[56] to include developing positive self-concept as a student of chemistry and minimizing the pressure to perform perfectly as to allow for mental space to think about chemistry concepts. Perhaps laboratory curricula that de-emphasize a sole focus on the outcome of the experiment and instead focus on giving students opportunities to make decisions and make sense of their data could increase students' cognitive and affective perceptions of their learning.

The results from this study also show that for many students, their expectations for learning in the laboratory shape how they perceive their experiences. Strike and Posner's statement in reference to conceptual change theory rings true here as well that "*Seeing is something we do with ideas as well as sense. We cannot see what we cannot conceive. Moreover, people who approach the world with different conceptions will see it differently*" (p 215).[57] For example, if students expect the laboratory to include affective experiences, but then are directed to focus only on cognitive experiences, this may minimize meaningful learning. If students expect the laboratory to focus only on concepts, but are then confronted with assessment that focuses only procedures, this, too, may minimize meaningful learning. Discussion of meaningful learning as the integration of cognitive and affective experiences could prove helpful for students (and instructors). Only when students *and* instructors recognize the learning experiences that contribute to meaningful learning in the laboratory can they be open to those experiences.

Implications for Research

This exploratory, descriptive study generated just as many questions as answers. This study used purposeful sampling among faculty to create a national sample that yielded diverse student responses for GC and OC. A criterion sampling protocol using type of university (Carnegie classification), type of laboratory pedagogy (problem based, cooperative learning, high inquiry), or diverse demographic information could yield new insights about students' perceptions of their learning. Additionally, these classifications could reveal new contributing factors to student learning. We chose to use the meaningful learning framework as a guide to make choices for the EFA and the cluster analysis. Thus, our results should be interpreted through this particular learning theory lens. Other theoretical lenses, particularly those focused on laboratory, might yield different insights, e.g., Kolb's experiential learning theory.[58]

The results showed students overall only encountered the experiences they expected at the beginning of the semester. How do students form these expectations for laboratory learning? What does the laboratory curriculum need to look like to specifically target students with low expectations for learning? The Change cluster raises additional questions for research. What elements of their experiences creates the gap

DOI: 10.1021/acs.jchemed.5b00538
J. Chem. Educ. XXXX, XXX, XXX−XXX

from their expectations? Are these students miscalibrated in some way about their learning? Does the laboratory curriculum not effectively meet their expectations? The Change cluster did not exist among the OC sample. Longitudinal studies are needed to investigate whether students in the Change cluster drop chemistry or if they recalibrate their expectations for learning based on their general chemistry experience.

Limitations

Interpretation of the data and findings presented to draw conclusions should be situated in the context to which the data were collected. The methodological choices we made impacted the data collection and analysis. Schools were recruited though an e-mail invitation to faculty who had previously expressed interest in research on laboratory learning and/or chemistry education research. Analysis of the pedagogy and curriculum in place at the data collection sites is still underway, so the current findings cannot be extrapolated to all laboratory environments. As the MLLI was administered online, some schools encouraged the students to complete the MLLI in the laboratory computer room while other students were given the opportunity to complete the survey at their convenience. Student participation at all schools was voluntary, though some schools did offer extra credit points if the class reached a certain threshold of participation. Attrition from the study does limit the conclusions that can be made from the data as we cannot make conclusions about those who did not participate at both the pre- and post-test administrations.

The methodological choices for the EFA and cluster analysis impact the results as well. The choice to explore the factors as potential mental frameworks for students' perceptions of learning, rather than to identify an internal factor structure for the MLLI, influenced how we interpreted the statistical output. We purposefully kept items that a strict interpretation of EFA output would have suggested we delete. Different clustering algorithms, clustering variables, distance measures, or linkage techniques might have produced different solutions. In addition, our decision to explore solutions greater than two clusters (even though two cluster solution had the largest distance on the dendograms and the highest VRC) yielded clusters with greater diversity that have the potential to lead to increased pedagogical utility by suggesting different interventions that would be more effective for Mid students than Change students.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available on the ACS Publications website at DOI: 10.1021/acs.jchemed.5b00538.

> Correlation matrices, additional inferential statistics, additional cluster visualizations, and interpretation of the exploratory factor analysis (PDF, DOCX)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: bretzsl@miamioh.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Volwiler Family Endowment to the Miami University Department of Chemistry & Biochemistry and by Grant No. 0733642 from the National Science Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors would like to thank Ellen Yezierski for her expert feedback on the research and statistics advice.

■ REFERENCES

(1) 2015 Guidelines for *Undergraduate Professional Education in Chemistry*, http://www.acs.org/content/dam/acsorg/about/governance/committees/training/2015-acs-guidelines-for-bachelors-degree-programs.pdf (accessed June 2015).

(2) Hilosky, A.; Sutman, F.; Schmuckler, J. Is Laboratory Based Instruction in Beginning College-Level Chemistry Worth the Effort and Expense? *J. Chem. Educ.* **1998**, *75*, 100−104.

(3) Elliott, M. J.; Stewart, K. K.; Lagowski, J. J. The Role of the Laboratory in Chemistry Instruction. *J. Chem. Educ.* **2008**, *85*, 145−149.

(4) Hofstein, A.; Lunetta, V. N. The Role of the Laboratory in Science Teaching: Neglected Aspects of Research. *Rev. Educ. Res.* **1982**, *52*, 201−17.

(5) Hofstein, A.; Lunetta, V. N. The Laboratory in Science Education: Foundations for the Twenty-first Century. *Sci. Educ.* **2004**, *88*, 28−54.

(6) Abraham, M. R. What Can Be Learning from Laboratory Activities? Revisiting 32 Years of Research. *J. Chem. Educ.* **2011**, *88*, 1020−1025.

(7) Lazarowitz, R.; Tamir, P. Research on Using Laboratory Instruction in Science. In *Handbook of Research in Science Teaching*; Gabel, D. L., Ed.; Macmillan: New York, NY, 1994, pp 94−130.

(8) Nakhleh, M. B.; Polles, J.; Malina, E. Learning Chemistry in a Laboratory Environment. In *Chemical Education: Towards Research-Based Practice*; Gilbert, J. K., De Jong, R., Justi, D. F., Treagust, J. H., Van Driel, J. H., Eds.; Kluwer: Dordrecht, The Netherlands, 2002; pp 69−94.

(9) Reid, N.; Shah, I. The Role of Laboratory Work in University Chemistry. *Chem. Educ. Res. Pract.* **2007**, *8*, 172−185.

(10) Sevian, H.; Fulmer, G. W. Student Outcomes from Innovations in Undergraduate Chemistry Laboratory Learning: A Review of Projects Funded by the US National Science Foundation Between 2000−2008. *Educ. Quim.* **2012**, *23*, 149−161.

(11) Jalil, P. A. A Procedural Problem in Laboratory Teaching: Experiment and Explain, or Vice-Versa? *J. Chem. Educ.* **2006**, *83*, 159−163.

(12) Golde, M. F.; McCreary, C. L.; Koeske, R. Peer Instruction in the General Chemistry Laboratory: Assessment of Student Learning. *J. Chem. Educ.* **2006**, *83*, 804−810.

(13) Woodfield, B. F.; Catlin, H. R.; Waddoups, G. L.; Moore, M. S.; Swan, R.; Allen, R.; Bodily, G. The Virtual ChemLab Project: A Realistic and Sophisticated Simulation of Inorganic Qualitative Analysis. *J. Chem. Educ.* **2004**, *81*, 1672−1678.

(14) Woodfield, B. F.; Andrus, M. B.; Anderson, T.; Miller, J.; Simmons, B.; Stanger, R. The Virtual ChemLab Project: A Realistics and Sophisticated Simulation of Organic Synthesis and Organic Qualitative Analysis. *J. Chem. Educ.* **2005**, *82*, 1728−1735.

(15) Hawkins, I.; Phelps, A. J. Virtual Laboratory vs. Traditional Laboratory: Which is More Effective for Teaching Electrochemistry? *Chem. Educ. Res. Pract.* **2013**, *14*, 516−523.

(16) Malina, E. G.; Nakleh, M. B. How Students Use Scientific Instruments To Create Understanding: CCD Spectrophotometers. *J. Chem. Educ.* **2003**, *80*, 691−698.

DOI: 10.1021/acs.jchemed.5b00538
J. Chem. Educ. XXXX, XXX, XXX−XXX

(17) Miller, L. S.; Nakleh, M. B.; Nash, J. J.; Meyer, J. A. Students' Attitudes Toward and Conceptual Understanding of Chemical Instrumentation. *J. Chem. Educ.* **2004**, *81*, 1801−1808.
(18) Tien, L.; Teichert, M. A.; Rickey, D. Effectiveness of a MORE Laboratory Module in Prompting Students to Revise their Molecular-Level Ideas about Solutions. *J. Chem. Educ.* **2007**, *84*, 175−181.
(19) Teichert, M. A.; Tien, L.; Anthony, S.; Rickey, D. Effects of Context on Students' Molecular-Level Ideas. *Inter. J. Sci. Teach.* **2008**, *30*, 1095−1114.
(20) Domin, D. S. Students' Perceptions of When Conceptual Development Occurs During Laboratory Instruction. *Chem. Educ. Res. Pract.* **2007**, *8*, 140−152.
(21) Cooper, M. M.; Kerns, T. S. Changing the Laboratory: Effects of a Laboratory Course on Students' Attitudes and Perceptions. *J. Chem. Educ.* **2006**, *83*, 1356−1351.
(22) Sandi-Urena, S.; Cooper, M. M.; Gatlin, T. A.; Bhattacharyya, G. Students' Experience in a General Chemistry Cooperative Problem Based Laboratory. *Chem. Educ. Res. Pract.* **2011**, *12*, 434−442.
(23) Sandi-Urena, S.; Cooper, M. M.; Stevens, R. Effect of Cooperative Problem-Based Lab Instruction on Metacognition and problem solving skills. *J. Chem. Educ.* **2012**, *89*, 700−706.
(24) Weaver, G. C.; Wink, D.; Varma-Nelson, P.; Lyle, F.; Morris, R.; Fornes, W.; Russell, C.; Boone, W. L. Developing a New Model to Provide First and Second-Year Undergraduate with Chemistry Research Experience: Early Findings of the Center for Authentic Science Practice in Education (CAPSiE). *Chem. Educator* **2006**, *11*, 125−129.
(25) Russell, C. B.; Weaver, G. C. A Comparative Study of Traditional, Inquiry-Based, and Research-Based Laboratory Curricula: Impacts on Understanding of the Nature of Science. *Chem. Educ. Res. Pract.* **2011**, *12*, 57−67.
(26) Szteinberg, G. A.; Weaver, G. C. Participants' Reflections Two and Three Years After An Introductory Chemistry Course-Embedded Research Experience. *Chem. Educ. Res. Pract.* **2013**, *14*, 23−35.
(27) Winkelmann, K.; Baloga, M.; Marcinkowski, T.; Guannoulis, C.; Anquandah, G.; Cohen, P. Improving Students' Inquiry Skills and Self-Efficacy through Research-Inspired Modules in the General Chemistry Laboratory. *J. Chem. Educ.* **2015**, *92*, 247−255.
(28) Novak, J. D. *Learning, Creating, and Using Knowledge*; Taylor & Francis Group: New York, NY, 2010.
(29) Novak, J. D. Human Constructivism: A Unification of Psychological and Epistemological Phenomena in Meaning Making. *Inter. J. Pers. Const. Psych.* **1993**, *6*, 167−193.
(30) Bretz, S. L. Novak's Theory of Education: Human Constructivism and Meaningful Learning. *J. Chem. Educ.* **2001**, *78*, 1107.
(31) National Research Council. *Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering*; National Academies Press: Washington, D.C., 2012.
(32) Ausubel, D. *Educational Psychology: A Cognitive View*; Holt, Rinehart, and Winston, Inc.: New York, NY, 1968.
(33) Galloway, K. R.; Bretz, S. L. Development of an Assessment Tool to Measure Students' Meaningful Learning in the Undergraduate Chemistry Laboratory. *J. Chem. Educ.* **2015**, *92*, 1149−1158.
(34) Fabrigar, L. R.; Wegener, D. T. *Exploratory Factor Analysis*; Oxford University Press: New York, NY, 2012.
(35) American Educational Research Association; American Psychological Association; National Council on Measurement in Education; Joint Committee on Standards for Educational and Psychological Testing (U.S.). *Standards for Educational and Psychological Testing*; American Educational Research Association: Washington, DC, 1999.
(36) Arjoon, J. A.; Xu, X.; Lewis, J. E. Understanding the State of the Art for Measurement in Chemistry Education Research: Examining the Psychometric Evidence. *J. Chem. Educ.* **2013**, *90*, 536−545.
(37) Bretz, S. L.; McClary, L. M. Students' Understanding of Acid Strength: How Meaningful Is Reliability When Measuring Alternative Conceptions? *J. Chem. Educ.* **2015**, *92* (2), 212−219.
(38) Cronbach, L. J. Coefficient Alpha and the Internal Structure of Tests. *Psychometrika.* **1951**, *16*, 297−334.
(39) Everitt, B. S.; Landau, S.; Leese, M.; Stahl, D. *Cluster Analysis*, 5th ed.; Wiley: Chichester, U.K., 2011.
(40) Galloway, K. R.; Bretz, S. L. Using Cluster Analysis to Characterize Meaningful Learning in a First-Year University Chemistry Laboratory Course. *Chem. Educ. Res. Pract.* **2015**, DOI: 10.1039/C5RP00077G.
(41) Mooi, E.; Marko, S. Cluster Analysis. In *A Concise Guide to Market Research*; Springer: Berlin, 2011; pp 237−284.
(42) Ward, J. H. Hierarchical Grouping to Optimize an Objective Function. *J. Am. Stat. Assoc.* **1963**, *58*, 236−244.
(43) Norusis, M. J. Cluster Analysis. In *Statistical Procedures Companion*; Prentice Hall: Upper Saddle River, NJ, 2011; pp 375−404.
(44) Burns, R. P.; Burns, R. Cluster Analysis. In *Business Research Methods and Statistics Using SPSS*; Sage Publications: London, 2009; pp 552−567.
(45) Caliński, T.; Harabasz, J. A Dendrite Method for Cluster Analysis. *Commun. Statist.* **1974**, *3*, 1−27.
(46) Milligan, G. W.; Cooper, M. C. An Examination of Procedures for Determining the Number of Clusters in a Data Set. *Psychometrika* **1985**, *50*, 159−179.
(47) Cohen, B. *Explaining Psychological Statistics*, 3rd ed.; Wiley: New York, 2008; pp 528−547.
(48) Cooper, M. M.; Sandi-Urena, S. Design and Validation of an Instrument to Assess Metacognitive Skillfullness in Chemistry Problem Solving. *J. Chem. Educ.* **2009**, *86*, 240−245.
(49) Marsh, H. W.; Scalas, L. F.; Nagengast, B. Longitudinal Tests for Competing Factor Solutions for the Rosenberg Self-Esteem Scale: Traits, Ephemeral Artifacts, and Stable Response Scales. *Psych. Assess.* **2010**, *22*, 366−381.
(50) Aldenderfer, M. S.; Blashfield, R. K. *Cluster Analysis*; Sage Publications: Beverly Hills, CA, 1984.
(51) Kaufman, L.; Rousseeuw, P. J.; *Finding Groups in Data: An Introduction to Cluster Analysis*; Wiley: New York, NY, 1990.
(52) Cohen, J. Eta-Squared and Partial Eta-Squared in Fixed Factor ANOVA Designs. *Educ. Psychol. Meas.* **1973**, *33*, 107−112.
(53) Levine, T. R.; Hullett, C. R. Eta Squared, Partial Eta Squared, and Misreporting of Effect Size in Communication Research. *Human Comm. Res.* **2002**, *28*, 612−625.
(54) Kirschner, P.; Meester, M.; Middelbeek, E.; Hermans, H. Agreement Between Student Expectations, Experiences, and Actual Objectives of Practicals in the Natural Sciences at the Open University of The Netherlands. *Int. J. Sci. Educ.* **1993**, *15*, 175−197.
(55) Könings, K. D.; Brand-Gruwel, S.; van Merriënboer, J. J. G; Broers, N. J. Does a New Learning Environment Come Up to Students' Expectations? A Longitudinal Study. *J. Educ. Psych.* **2008**, *100*, 535−548.
(56) Bretz, S. L.; Fay, M.; Bruck, L. B.; Towns, M. H. What Faculty Interviews Reveal About Meaningful Learning in the Undergraduate Chemistry Laboratory. *J. Chem. Educ.* **2013**, *90*, 281−288.
(57) Strike, K. A.; Posner, G. K. A Conceptual Change View of Learning and Understanding. In *Cognitive Structure and Conceptual Change*; West, L. H. T., Pines, A. L., Eds.; Academic Press: Orlando, FL., 1985.
(58) Towns, M. H. Kolb for Chemists: David A. Kolb and Experiential Learning Theory. *J. Chem. Educ.* **2001**, *78*, 1107.

DOI: 10.1021/acs.jchemed.5b00538
J. Chem. Educ. XXXX, XXX, XXX−XXX