JOURNAL OF CHEMICAL EDUCATION

pubs.acs.org/jchemeduc Article

Reformed Experimental Activities (REActivities): Gauging the Fidelity of Implementation in a Reformed Undergraduate Organic Chemistry Laboratory

Christina Goudreau Collison,* Dina L. Newman, Julia Qingli Biehler, Micaela Nelson, Paige O'Brien Daly, Matthew Jackson, Cameron Isaac, Madeline Tebrugge, Jason Anderson, Brian Edelbach, Douglas Tusch, and Jeremy Alan Cody

Cite This: J. Chem. Educ. 2022, 99, 2032−2043 Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Reformed Experimental Activities (REActivities) is an innovative approach to the delivery of the traditional material in an undergraduate organic chemistry laboratory. To better understand the fidelity of implementing this pedagogy and what effects the framework of REActivities has on student−instructor interactions, an observational protocol study was employed. This paper also describes the Evaluation of Lab Instructor Time and Engagement (ELITE) observational instrument developed to evaluate instructional behaviors in a lab setting. The instrument was used in first semester undergraduate organic chemistry laboratories across seven universities to measure the robustness of the guided-inquiry materials when implemented and the nature of the instructor's interactions. The ELITE data was analyzed for laboratories delivered using REActivities and compared with expert delivery as well as with traditional expository methods. The data revealed that instructor behaviors when using REActivities were consistent and comparable and could be distinguished from traditional lab deliveries. The nature of the instructor's behaviors also showed a remarkably consistent trend for coded conceptual-based discussion when REActivities was employed, in contrast to the near absence of similar interactions when expository methods were employed for comparable laboratories.

KEYWORDS: *Laboratory Instruction, Chemical Education Research, Organic Chemistry, Second-Year Undergraduate, Inquiry-Based/Discovery Learning, Student-Centered, Learning*

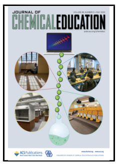
FEATURE: Chemical Education Research

INTRODUCTION

Reformed Experimental Activities (REActivities) was recently described as a scaffolded lab curriculum for undergraduate organic chemistry laboratories that allows students to practice inquiry-learning.[1−3] The REActivities approach acknowledges that organic chemistry is complex and requires multiple levels to understand a concept, and learners are limited in the amount of new information they can learn at any one time.[4−9] Previous research has shown that the materials[10] and the instructor influence the level at which students discuss a concept.[11−13] Both can help influence the discussion to assist in building connections, making multilevel discussions about concepts the norm. This is important because if we want enhanced student engagement and better connectivity of course material between the lecture and laboratory, we need to develop materials and prompts that elicit evidence of both.[14−19]

One trademark component of REActivities is the lack of a prelab lecture or assignment where the conceptual underpinnings of the lab are typically explained. Instead, REActivities strategically incorporates those same conceptual underpinnings at the appropriate moments *during* the lab through guided questions.[1,20−23] As such, students learning with REActivities begin the hands-on activities of the lab immediately upon arrival (the self-start). As students use the formalized practice time for techniques and then conduct the lab experiments, they are prompted to check in with their lab partner as well as with their instructor to discuss their answers to concept questions before moving on. Because inquiry-based learning yields superior results in student attitude, laboratory performance, and future problem-solving abilities over traditional didactic learning techniques,[24−34] REActivities is designed to provide opportunities for students during the lab period to think critically

Received: September 21, 2021
Revised: March 29, 2022
Published: April 14, 2022



https://doi.org/10.1021/acs.jchemed.1c00998
J. Chem. Educ. 2022, 99, 2032−2043

Table 1. Participating Institutions and Their Classifications Where Undergraduate Organic Chemistry Lab Sessions Were Video Recorded for ELITE Coding[a]

	instructor						
university	REActivities	expository	status	university type	program type	enrollment	students in section
A	1[b]		private	research university	4-year	16 300	22
B	2		public	community college	2-year	11 600	15
C	3		public	primarily undergraduate	4-year	3760	13
D	4		private	comprehensive	4-year	1500	12
E	5	6	public	research university	4-year	10 600	18
F		7	private	research university	4-year	6500	20
G	8		private	liberal arts	4-year	1600	14

[a] All data taken from Carnegie Classifications and National Center for Education Statistics. [b] The instructor who designed and implemented REActivities at the home institution (A).

about chemistry concepts and ask questions that support discovery.

Before the effectiveness of our reformed instructional practices on student learning could be studied, we needed to understand the fidelity of implementing REActivities by various instructors across institutions.[35−38] Fidelity studies establish that the instructional practice is delivered as the design was intended by monitoring the implementation and comparing it to a control teaching method. It has been shown that instructors adopting a new instructional practice often inadvertently alter or omit critical components of the pedagogy,[39−42] but few studies calibrate the implementation of instructional practices prior to the measurement of student learning outcomes. The validity of future studies on student outcomes when using REActivities will rely largely on understanding how and if the instructional practice is broadly implemented with integrity. This integrity is very important or else any positive learning gains reported could be attributed to factors other than adoption of REActivities.

Fidelity of Implementation

Stains and Vickrey developed a framework and methodological approach for characterizing the fidelity of implementation (FOI) of an instructional practice[35,39,43−45] The approach involves first identifying both the structural components and instructional components deemed critical to successful delivery of the pedagogy. The structural components for REActivities would comprise the written workbook materials and instructor training practices required when delivering REActivities. The instructional components specifically would relate to the behaviors of both the instructor and the student during the delivery of REActivities laboratories.[43,44,46]

As mentioned, REActivities incorporates guided questions and prompts during the relevant hands-on lab activities. The workbook and its guided script serve as an important structural component that helps inform the students and instructor what they should do during implementation. REActivities is not expected to rely on seasoned instructors and was originally created with the novice graduate teaching assistant (GTA) in mind. As such, no high level of pedagogical knowledge is expected to be necessary for successful implementation of REActivities. It is anticipated that the REActivities materials will guide both the student and instructor through the lab exercise, outlining each of their activities and roles without a large degree of training. A second structural component of REActivities is the self-start, which works in concert with the third structural component, the near absence of lecturing by the instructor. The last structural component requires that the instructor be present and not be engaged in personal activities. The term "actively inactive" was coined by one of the adopters of REActivities. Instructors are encouraged to actively monitor the lab by walking around to physically signal to the students that the instructor is available and receptive to questions.

The instructional components deemed critical for successful delivery of REActivities comprises the student−student and student−instructor interactions and the nature of these interactions as set forth by the REActivities materials. Thus, a strategy to characterize these components when REActivities laboratories are delivered was developed in the form of an adapted observational protocol instrument. No perfect observational instrument was available to capture all structural and instructional components for our study. As such, in order to successfully analyze instructor behaviors and track FOI, a new observational instrument needed to be designed. Inspired by the Real-Time Instructor Observation Tool (RIOT),[47] and the Learning Observational Protocol for Undergraduate STEM (LOPUS),[37] the Evaluation of Lab Instructor Time and Engagement (ELITE) Instrument was developed to meet the needs of this study. The ELITE instrument combines the continuous instructor−student behavior coding protocol of RIOT with the instructor engagement codes found in the segmented LOPUS protocol.

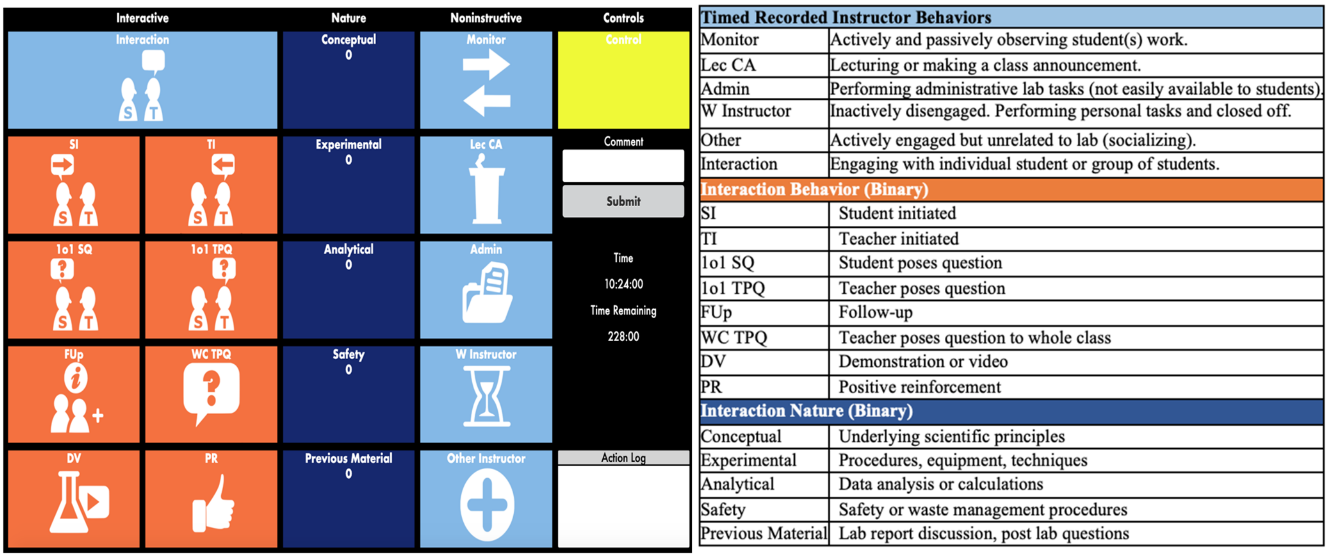
To conduct this research, a multi-institution study was organized with five universities willing to participate alongside the home institution in full adoption of REActivities for the first semester of an organic chemistry lab and one university willing to serve as an expository lab control. Using multiple instructors across institutions added variability to the study. It was important that each participating instructor have some training to deliver the REActivities pedagogy and that they be consistent in their delivery. This study would thus be critical to measure the degree of FOI across REActivities laboratories in order for our subsequent studies on student learning outcomes to be valid.

■ RESEARCH QUESTION

This study focuses on FOI of REActivities in the first semester of undergraduate organic chemistry laboratories taught by different instructors across institutions. Specifically, we sought to answer the following question:

To what extent does the framework of REActivities impact the fidelity of implementation as it relates to the intended structural and instructional components?

To do so, an observational protocol was designed to capture instructional behaviors during a lab period, and our study also sought to validate that our observational instrument was sensitive enough to capture the hypothesized critical components of the curriculum.



Timed Recorded Instructor Behaviors	
Monitor	Actively and passively observing student(s) work.
Lec CA	Lecturing or making a class announcement.
Admin	Performing administrative lab tasks (not easily available to students).
W Instructor	Inactively disengaged. Performing personal tasks and closed off.
Other	Actively engaged but unrelated to lab (socializing).
Interaction	Engaging with individual student or group of students.
Interaction Behavior (Binary)	
SI	Student initiated
TI	Teacher initiated
1o1 SQ	Student poses question
1o1 TPQ	Teacher poses question
FUp	Follow-up
WC TPQ	Teacher poses question to whole class
DV	Demonstration or video
PR	Positive reinforcement
Interaction Nature (Binary)	
Conceptual	Underlying scientific principles
Experimental	Procedures, equipment, techniques
Analytical	Data analysis or calculations
Safety	Safety or waste management procedures
Previous Material	Lab report discussion, post lab questions

Figure 1. ELITE instrument interface displaying coder buttons, a key detailing the acronyms used, and brief description of code use.

■ METHODS

Data Set

The laboratories studied were first semester undergraduate organic chemistry laboratories. Data was collected from a diverse group of institutions (Table 1) varying from two-year to four-year schools, community college to research institutions, comprehensive universities to primary undergraduate institutions, and size from 1500 to 16 300 students. There were varying degrees of experience among the instructors observed, but most instructors were faculty and one was a GTA. Three of the eight instructors in the study were females. Instructor 1 at University A is the original developer of REActivities.

Since the scope of our work in the future would extend beyond FOI, universities within a 3-h driving radius were targeted. Prior to 2019, proximity was considered important so that participating institutions would be more accessible when conducting interviews with both students and instructors. Of the 10 institutions asked to adopt REActivities, five agreed to adopt and participate (universities B−E and G). The instructors at University E were not in agreement about adoption so it was allowed that each instructor could teach the material using their method of choice and both permitted their laboratories be recorded. Institutions that declined were not interested in changing their organic chemistry lab practices and materials and were typically larger institutions utilizing GTAs as instructors. University F was one such institution that declined to adopt REActivities given time and resources but did agree to have their laboratories recorded.

Instructor Training

In order to obtain a level of consistency when delivering the REActivities pedagogy, instructor training was incorporated prior to the study. The instructors from participating universities attended a one-day workshop during the summer prior to adoption. The workshop included a mock REActivities lab where the instructors acted as the students and were led through the following key principles to deliver a REActivities lab:

- **Self-start**: Instructors were told not to give a prelab lecture. Prelab assignments were discouraged but information about the lab could be posted. They were instructed to simply encourage immediate activity once the students arrive to lab.
- **Flexibility**: Generally, students complete the entirety of the material for a given lab in the allotted time period. Sometimes students do not have enough time to complete all activities for a lab due to time.[1] For these scenarios, participating instructors were shown multiple ways to stay flexible with assignments in order to keep anxious students from rushing but without falling behind. Additionally, instructors were encouraged to be flexible with lab report write-up due dates but more stringent on workbook or postlab question due dates. This rhythm is more in line with research activity and changes the rush-to-get-out mindset, thereby slowing the pace down to create more time and space for discussion.
- **Presence**: Instructors were trained on how to physically present themselves as available to the students. This meant that they needed to be actively present at all times. It is OK if instructors are busy with something lab-related or quietly walking around monitoring the room. It is the wrong assumption that since the material is self-guided, an instructor would have more time to work on personal tasks or be on their phone. Instructors of the pedagogy often report that they are more tired after delivering a REActivities lab than a traditional lab approach, due to the constant walking around and engagement.

Data Collection

First semester undergraduate organic chemistry teaching laboratories from all of the institutions were video recorded and coded for the instructor's behavior during the lab and the nature of the interaction they had when engaging with students. Two institutions had instructors delivering the material using traditional methods. These traditional laboratories served as a control group for comparison. Since first semester organic chemistry laboratories focus on techniques, it was easy to map REActivities laboratories with their expository counterpart.

Human Subjects Institutional Review Board (HSIRB) approval was obtained for the study (46.101 (b)[1]) and all participants, including instructors and students, provided consent to be video and audio recorded. After consent, the

Table 2. Structural and Instructional Components Deemed Critical to Delivery of REActivities Mapped to the Assessment Ability of the Video Recording Validation and the Evaluation of Lab Instructor Time and Engagement (ELITE) Observation Protocol

	Yes/No Validation	Interactive Behaviors	Lec/CA	Interaction	Admin	Other	Monitor	Nature
Structural Components								
Workbook utilization	✓							
Student self-start	✓		✓					
Absence of Lecture			✓					
Actively-inactive instructor		✓	✓		✓	✓	✓	
Instructional Components								
Student-student engagement	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
Instructor-student engagement		✓		✓				
Nature of student-student engagement	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
Nature of instructor-student engagement								✓

undergraduate organic chemistry laboratories of participating institutions were video recorded using a GoPro, while the instructor attached a Bluetooth microphone onto their lab coat for audio recording. Having the Bluetooth microphone attached to the instructor's lab coat allowed for focused capture of each conversation between the instructor and a student. As with any technology, technical difficulties occurred at times, so some of the data collected was not usable. The data was considered unusable if the audio recording or video recording cut out, since both audio and video are required to fully code the lab instructor's interactions and conversations. Due to these technical limitations, the usable data acquired was naturally random.

Development of the ELITE Observational Instrument

In order to successfully analyze instructor behaviors and track FOI, a new observational instrument was designed. The Evaluation of Lab Instructor Time and Engagement (ELITE) protocol uses a majority of the codes found in the LOPUS protocol[37] and combines them with the continuous coding practice found in the RIOT protocol.[47] To assist coders in recording and mapping instructor behaviors, the ELITE instrument was created using the Generalized Observation and Reflection Platform (GORP) interface.[48] This interface allowed for facile and descriptive coding for the nature of each interaction and most importantly, the duration of each behavior, thus accounting for 100% of the instructor's time during a lab period.

Coding

To establish reliability for ELITE coding, three trained student observers simultaneously scored each laboratory recording in intervals of 20 min. The ELITE interface (Figure 1) uses a color system to easily code for instructor behavior types and the nature of the interactions. When beginning to observe a recorded laboratory delivery, three trained coders initiate the coding by simultaneously hitting the yellow button, which is labeled "Control" in order to align their codes to a time stamp. This button remains on throughout the entirety of a lab, allowing numerous other behaviors to also be coded. When more than half of the students have finished and left the laboratory, signaling that the lab is over, the coders click off this control button, which finishes a coding session. Discrepancies in coding were most often a mistakenly pressed button and were resolved easily by using the time stamp on the ELITE instrument to efficiently find and rewatch that portion of the recording.

The light blue buttons code for the instructor's behaviors (Interaction, Monitor, Lecture/Class Announcement (Lec/CA), Administration (Admin), Waiting, and Other). When a light blue button is pressed, a timer begins for that given behavior. Thus, these buttons time the duration of an instructor's activity resulting in one of the light blue behavior buttons being active at all times during the lab period. The following behavior buttons are used as follows:

- **Interaction**: Reserved strictly for instructor−student interactions when the topic of engagement concerns the material for the lab.
- **Monitor**: Used when the instructor is walking around the lab and overseeing the activities by the students but not engaging in any conversations.
- **Lec/CA**: (Lecture/Class Announcement) Indicates that the instructor is lecturing to the lab as a whole or making a class-wide announcement.
- **Admin**: Used when the instructor is actively doing something related to the lab but not necessarily interacting with the students. For example, the instructor may be emptying a waste container, getting more materials for the lab, or passing out last week's lab reports.
- **Other**: Used when the instructor is engaged with a student or lab visitor but discussing topics that are not about the lab content. Often these interactions are more social in nature.
- **Waiting**: (W Instructor) Indicates that the instructor is not interacting with the student and is busy with personal activities such as working on their phone or computer, grading papers, etc.

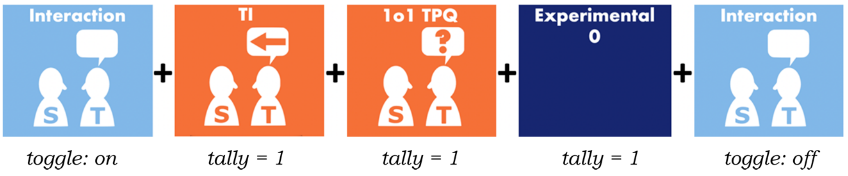


Figure 2. An example of the ELITE codes utilized for scenario 1 using both the timed response behavior buttons (light blue), the instance codes for type of activity (orange), and the instance code for nature of the interaction. The coder would activate: Interaction < TI < 1o1TPQ < Experimental < Interaction.

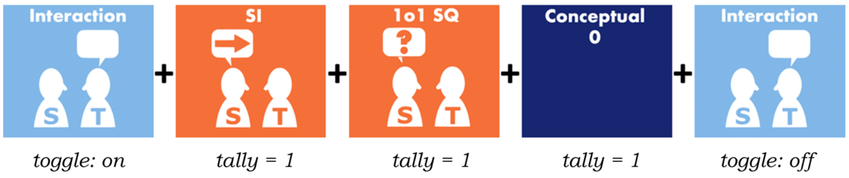


Figure 3. An example of the ELITE codes utilized for scenario 2 using both the timed response behavior buttons (light blue), the instance codes for type of activity (orange), and the instance code for nature of the interaction. The coder would activate: Interaction < SI < 1o1SQ < Conceptual < Interaction.

The orange buttons code for the details of the interactive behavior (student-initiated, teacher-initiated, student-posed question, instructor follow-up, teacher posed question, etc.) and are not timed; instead, they tally the number of instances for the given type of behavior. These orange interaction behavior codes were included because they were also present in the original LOPUS instrument and were predicted to be valuable when coding for our critical components as it related to instructor−student engagement. In the end, these orange interactive behavior codes yielded no informative trends for this study. Thus, despite using them, their analysis will not be discussed further.

When a light blue interaction button is activated, it is required that a dark blue button is also coded. The dark blue buttons code for the nature of the interaction and also tally instances instead of duration. The most commonly coded interaction types during a lab period are experimental and conceptual. Experimental interactions are procedural- or protocol-based questions or comments, while conceptual interactions include a more in-depth question or comment about the underlying scientific and molecular principles related to the protocol. The three less-observed interaction types include analytical, safety, and previous material. An analytical interaction involves questions and discussion about calculations within the procedure; safety strictly pertains to comments regarding waste, and other safety protocols; and previous material codes are used when students need assistance reviewing experimental techniques.

How the design of the ELITE instrument would capture the structural and instructional components deemed critical to the REActivities pedagogy was correlated (Table 2). The fact that each lab period needed to be recorded for coding allowed for some structural components to be measured and validated as either present or not present (Yes/No). For example, watching the lab recording clearly showed students using the easily recognizable lab workbook and coders could make note of this. Coders could also watch the video and validate that the students arrived and immediately got to work at their bench or station (self-start). Absence of a prelab lecture was also noted, but the validation that lecture remained absent throughout the lab and what the exact behavior map was for the instructor would require the ELITE instrument. Furthermore, the actively inactive behavior of the instructor could be tracked using the three ELITE codes for "monitor", "admin", and "other". These three codes ensure that the instructor is open to student engagement. In contrast, an instructor who is busy with personal activities (waiting code) signals to the student that they are less approachable.

Instructional components that involved the instructor were measurable by the ELITE instrument. This was due to the fact that the instructor was the only person wearing the microphone. Coded conversations always involved the instructor and were not sensitive to picking up peripheral student−student engagements. As such, the student−student interactions and the nature of their engagements could not be measured by the ELITE instrument and would also require a different recording and microphone setup.

What follows are actual scenarios from the recordings demonstrating coding instances.

Scenario 1—Instructor ends monitoring and approaches a student during the distillation lab (Figure 2):

Instructor: Did you do a thermometer check? Did you see the image of where your thermometer should be?

Student consults workbook and compares image to their setup.

Student: It needs to be a little lower.

Instructor: Yeah.

Scenario 2—Instructor is stopped by two students during the thin layer chromatography lab (Figure 3):

Student 1: Right, because this does not have an OH... these have OH's.

Student 2: Yeah, that is what I'm thinking, so I think it is 1, 2, and then I think it is 3, 4, 5 (as she points to individual molecules in the workbook trying to rank their polarity).

Student 1: So, we think this is a 2?

Student 2: (Steps back and asks instructor) Am I wrong?

Instructor: Hey, you know the best part of this? You can guess, and you come up with some reasoning, and then you can test it. And then if it is right, it is right, and if it is wrong, then the question is what did we miss?

Reliability, Validation, and Data Analysis

To establish inter-rater reliability for ELITE coding, three trained student observers simultaneously scored each laboratory recording in intervals of 20 min. Between each segment, the coders discussed their scoring iteratively until overall scoring agreement was reached. The lab was coded continuously for the

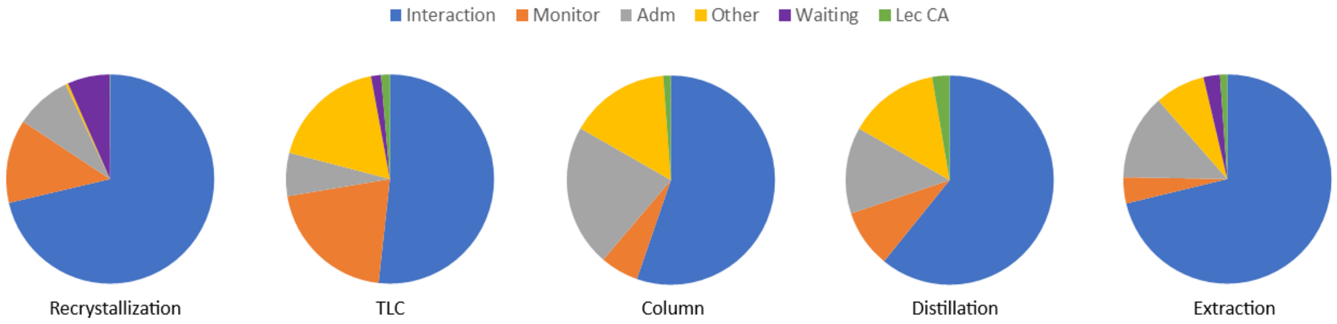


Figure 4. Behavioral time map of the expert instructor (A1) using REActivities for the duration of a four-hour lab period with 22 students in the section. The lab techniques are represented from left to right in the order they were delivered in the first semester of an organic chemistry lab.

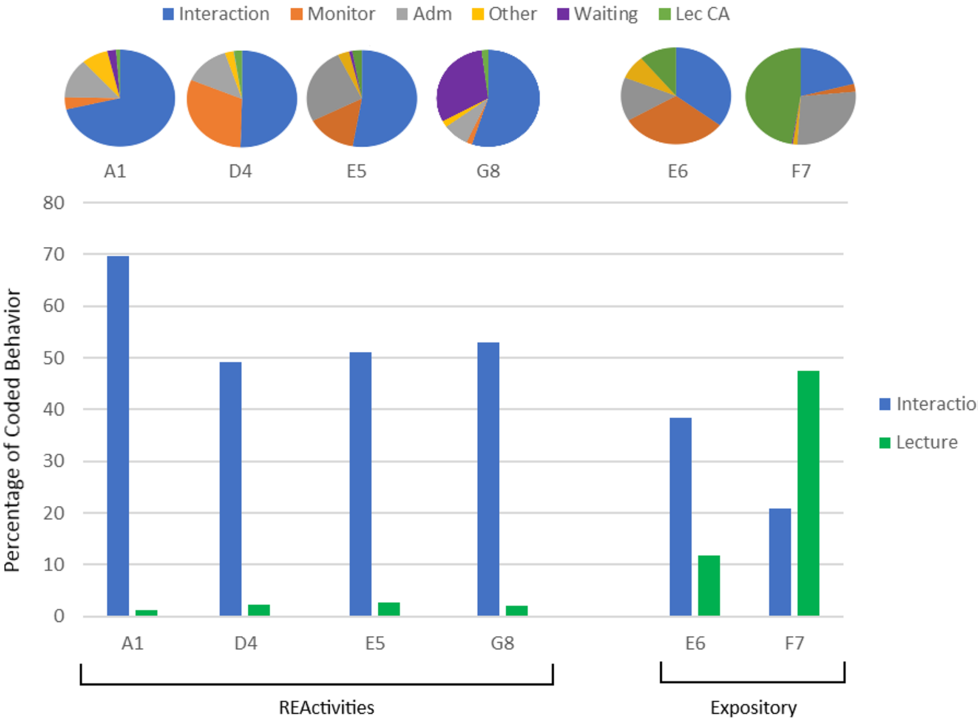


Figure 5. Behavioral time maps of the expert instructor (A1), three different instructors at different institutions (D4, E5, G8) using REActivities, and two different instructors using expository lab delivery (E6, F7) for a liquid−liquid extraction lab. Comparative timed behaviors indicating the percentage of time the instructor was observed either lecturing to the class or interacting with a student during the extraction lab is below its representative pie chart.

duration of the entire lab session, allowing for coders to quantify and account for 100% of the instructor's behaviors during a lab period. The data was broken down by each continuous behavior or tallied code. The validation process for continuous behavior data compares the duration of each coded interaction. If the duration of each coder's mapped interaction was within a 10 s time frame, all three coders' data were accepted and averaged together. If the duration of one coded interaction was an outlier compared with the other two coders for the same instance, the coders could easily rewatch segments of the recording using the time stamps to reach an overall agreement. Omitted instances were rare (<5%) and would typically comprise a short code investigated to be an error by one coder. These mistake instances were omitted and not included in the average. For the validation of tallied instances (orange and dark blue codes), the start time of the code was evaluated across all three coders. In order to accept a tallied behavior, all coded instances needed to be mapped with one another; otherwise, the outlying tallied instance was omitted. Again, the use of three coders and time stamped videos ensured that omissions for tallied instances were also rare (<2%).

■ RESULTS

The data from the ELITE coding was analyzed by first evaluating the laboratories of an expert user of REActivities, namely the author of the method. The data acquired by analysis of the expert user was considered the baseline for comparison. Once confirmed that the structural and instructional components of the delivery by the expert instructor were successfully captured by the ELITE instrument, the baseline behavior maps were then compared to those of other instructors across different institutions adopting REActivities. Lastly, this data was compared with the ELITE codes for faculty delivering the same material using expository methods.

Establishing a REActivities Baseline

The structural and instructional components must first be established as reliable and valid by evaluating a highly trained instructor of the reformed teaching practice.[43,46,49,50] In our case, this expert was the creator of the REActivities pedagogy and the laboratories recorded were captured in the third year of REActivities implementation at the home institution. It was important that all lab recordings for each instructor were not representative of their first semester of adoption. The ELITE baseline behavior maps were recorded for that same instructor over the course of one semester for a suite of first semester technique laboratories using REActivities (Figure 4). Notice that the two expected structural components for REActivities laboratories delivery when implemented correctly were captured in the behavior maps for the expert instructor. First, the near absence of Lecture/CA coding ensures that the self-start principle was adhered to and that a formal prelab lecture was abandoned. Second, the instructor spent a good amount of time engaging with students about the lab and making themselves available to the students. When not engaged with students on relevant lab content, the instructor was also largely chatting with students socially (other), quietly monitoring the lab by walking around (monitor), or seeing to pertinent lab-related tasks (admin). A very small percentage of time was coded as waiting indicating that this instructor made themselves available to the students and was rarely focused on personal activities.

An important instructional component captured by the instrument was instructor−student engagement. A healthy degree of instructor−student interactions was observed across all lab experiments for instructor A1. The nature of the interactions would be analyzed separately. At this point, the baseline established an estimation for the appropriate amount of instructor−student engagement given a lab size of 22 students working in pairs. If the REActivities framework prompts student−instructor engagement, then having more students should correlate to a larger time engaging with students. The meaning of how much time coded for admin, monitoring, or other were considered unique to that instructor's teaching style and did not reveal emerging trends. Thus, FOI across participating institutions can be ensured if there is a consistent lack of Lecture/CA and an appropriate amount of instructor interaction durations proportional to the number of students in a lab section. In order to gauge how these same laboratories were delivered at other institutions using REActivities the data from instructor A1 was used as the baseline for comparison in subsequent analyses.

Fidelity of Implementing REActivities Across Institutions

Using the data from Figure 4 as a baseline, FOI was evaluated when REActivities was utilized at other institutions with a deliberate focus on the instructor−student interaction code and the lecture code. Each participating institution's lab sections varied in size, lab period duration (3- and 4-h laboratories), instructor, and student composition (majors vs nonmajors). One constant was that the technique laboratories were delivered in the same sequence across all institutions. An example comparison involves the liquid−liquid extraction lab since we had usable data for a variety of instructors using REActivities as well as two expository lab sections (Figure 5).

There was a definitive trend in the instructor−student interactions when REActivities was utilized typically exceeding 40% and minimal occurrence of lecture-type behavior. In contrast, laboratories coded when expository methods were employed (Figure 5, E6 and F7) showed less instructor−student engagement time and a marked increase in lecture-type behavior. Of note, instructors E5 and E6 were from the same institution but chose to deliver the material differently during the allotted 3-h lab period. Instructor E5 implemented REActivities for liquid−liquid extraction while instructor E6 delivered the extraction content using a traditional approach. The larger amount of Lecture/CA coding for instructor E6 was due to 20 min of lecturing to the class at the start of the 3-h lab period. Instructor F7 had a 1.5 h prelab lecture followed by a lab period that ran just under three hours accounting for the larger lecture code. Again, the extraction laboratories using REActivities show minimal Lecture/CA coding and an appropriate amount of instructor interaction time consistent with the baseline ELITE data for instructor A1. This comparative analysis of the behavioral trends of REActivities against expository lab instruction reassured us that ELITE was sensitive enough to distinguish such differences. The next question was whether these trends were observed across all technique laboratories recorded and coded using ELITE.

Of the more than 40 REActivities laboratories recorded, a total of 25 lab sessions were usable for coding instructor behaviors, and of the 10 expository lab sessions, only 5 were usable for coding instructor behaviors. The five baseline laboratories of the expert instructor (A1) are included in the 25 REActivities laboratories. Averaged percentages for instructor behaviors for all coded laboratories when instructors use REActivities versus when instructors use expository methods could then be analyzed and compared (Figure 6). Of note is the diminished amount of Lecture/CA behavior and a healthy amount of interaction time for instructors using REActivities. The large range observed for interaction time mostly correlated to the number of students in the lab section. The two outlying data points for REActivities showing more than average lecture behavior was for the same instructor who for two laboratories early in the semester went off script and gave some prelab lecture to the students. Despite these two outliers, on average the REActivities materials were robustly holding together the structure of the lab course.

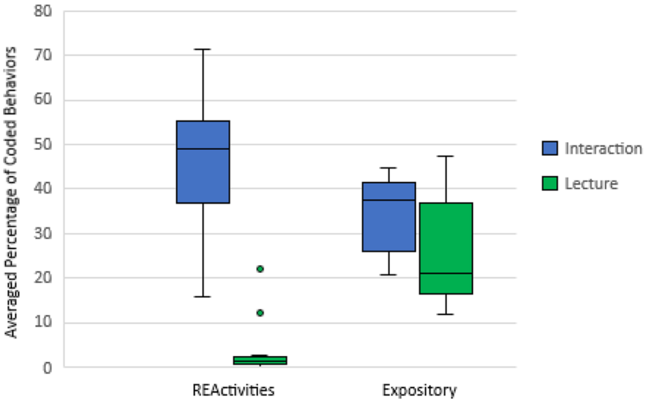


Figure 6. Box plots of coded behaviors for all participating instructors for all technique laboratories using REActivities ($n = 25$) in comparison to the coded behaviors of instructors delivering the same material in an expository fashion ($n = 5$).

For instructors using expository methods, a significantly larger average of coded behaviors for lecture was observed ($p < 0.0001$) when REActivities was compared with expository delivery, and a somewhat significant difference in averages was observed for coded student−instructor interactions ($p = 0.0946$) when REActivities was compared with expository delivery (Table 3). In recognition of the varying sample size for

Table 3. Statistical Results for Paired Comparisons between REActivities and Expository Labs for the Average Percentage of Coded Behaviors

	REActivities vs expository interaction $n = 25$	REActivities vs expository lecture $n = 5$
t test	$p = 0.0946$	$p < 0.0001$
Hedge's *g*	0.8475	3.041

each group, effect size was analyzed using Hedge's *g* calculations and determined to be very significantly different ($g = 3.041$) for lecture behavior between the two delivery methods.[51,52] Note that in all REActivities laboratories, the largest fraction of instructor time was spent directly interacting with students, as opposed to expository laboratories, where the instructor spent less of their time interacting with students and much more time talking at them.

Comparing the Nature of Interactions

The ELITE instructor behavior maps indicate a healthy amount of instructor interactions with students for REActivities and some student−instructor engagement in expository laboratories. The most impactful data was observed when the instructional component involving the nature of the instructor−student interactions was evaluated. Of the 25 REActivities lab videos usable for coding behavior, only 14 REActivities laboratories were usable for both video and audio quality in order to confidently code the nature of each instructor−student conversation. All five expository recordings were usable for both audio and visual quality.

The liquid−liquid extraction REActivities lab is again spotlighted to show data parsed out by institution to compare the expert user (A1) with three different instructors across universities to correlate similarities to instructors adopting REActivities (Figure 7). Although there were five nature categories, greater than 95% of the coded instances were either experimental or conceptual. Thus, comparisons were made focusing only on these two nature codes. The expert user of REActivities (A1) showed that when they engaged with students during the lab, a consistent degree of conceptually coded interactions was observed and this general trend extended to the other three adopters of REActivities (D4, E5, and G8). As expected, each lab showed a high number of experimental instances when compared to conceptual instances since the organic chemistry lab manipulations performed are new experiences for the students. However, it was surprising that for both instructors using expository methods (E6 and F7), a near absence of conceptual-based interactions was observed. Again, it is interesting to compare instructor E5 and E6 since they are from the same institution but chose to deliver the content using different formats. Instructor E5 used REActivities while instructor E6 used a traditional approach. Despite the behavioral map for these same instructors in Figure 5 showing comparable interaction times with the students, a noticeable difference in the nature of the interactions was observed (Figure 7). It is important to recognize that the nature codes were being analyzed during the instructor−student interaction time and does not represent what was presented during an instructor's lecture/CA time.

Prompted by the consistency of coded conceptual instances when REActivities is used during the extraction lab, data from all 14 usable REActivities laboratories were analyzed together to investigate if similar ratios for experimental and conceptual instances are observed across each lab technique activity (Figure 8). Regardless of instructor or institution, the data suggests that REActivities ensures a consistent ratio of conceptual to experimental topics during instructor−student interactions, indicating that instructors and students were engaging in discussion about underlying scientific principles rather than only troubleshooting a lab protocol. It was found that on average 27% of all coded interactions were conceptual when REActivities was utilized. The consistency of conceptual instances across each lab experiment at different institutions additionally supported the measurement for fidelity of implementation. When conceptual conversations and experimental conversations were compared, a statistically significant difference was observed ($p < 0.001$) between delivery methods (Table 4). Due to the difference in pool size between REActivities and expository data sets, a Hedge's *g* effect size was again utilized showing a very significant difference between experimental conversations ($g = 2.785$) and conceptual conversations ($g = 3.041$). Since REActivities laboratories

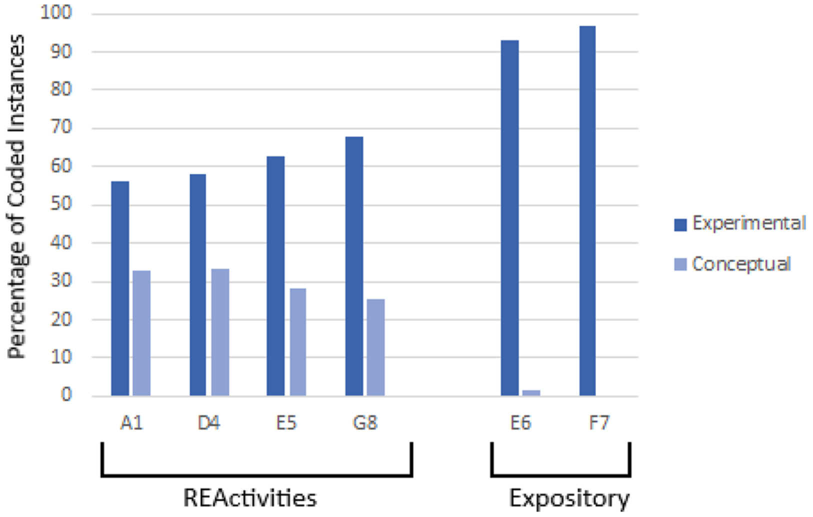


Figure 7. Experimental vs conceptual interactions between student and instructor across multiple institutions for a liquid−liquid extraction lab where A1 is an expert user, D4, E5, and G8 are REActivities adopters, and E6−F7 use an expository lab delivery.

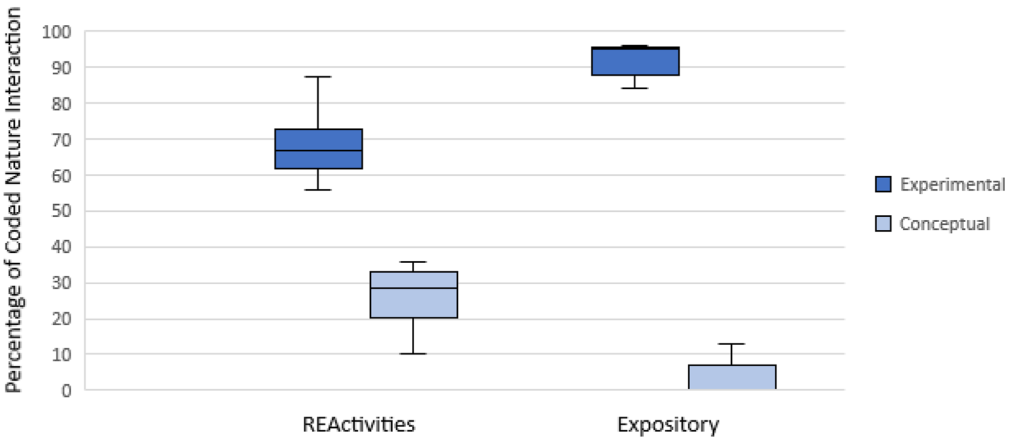


Figure 8. Averaged experimental vs conceptual interactions for five REActivities experiments across participating institutions (n = number of different instructors at different institutions coded).

Table 4. Statistical Analysis of Comparison between REActivities and Expository Delivery for Nature of Coded Instances

	REActivities vs expository	REActivities vs expository
	experimental $n = 14$	conceptual $n = 5$
t test	$p < 0.0001$	$p < 0.0001$
Hedge's *g*	2.785	3.041

prompt students to discuss how they correlate macroscopic activities to the conceptual underpinnings of the chemistry with their instructor, the presence of such conceptual interactions means that the guided inquiry nature of REActivities is being captured. This is key because students are not only asking about the lab's technical aspects, but also engaging in discussions with the instructor involving the underlying principles behind the laboratory experiment.

Relationship Between Engagement and Nature of Interactions

Despite differences in instructors' teaching style or experience, there is still evidence of FOI when delivering REActivities. A few anomaly instructor behavior maps were observed. These outliers showed little to no Lec/CA but had lower than expected instructor−student interaction time. For example, Figure 9 shows two very different behavior maps for two instructors using REActivities during a recrystallization lab. The differences in the observed interaction times could not be justified by the number of students in the lab since there were approximately 13 students in each section. The differences were thus attributed to teaching styles, given that instructor C3 was observed walking around the lab more and quietly monitoring the students, while instructor G8 was often performing administrative tasks at the front of the lab. Regardless of the instructor's teaching style, the ratio of experimental and conceptual interactions was consistent, indicating the importance of examining the nature of the

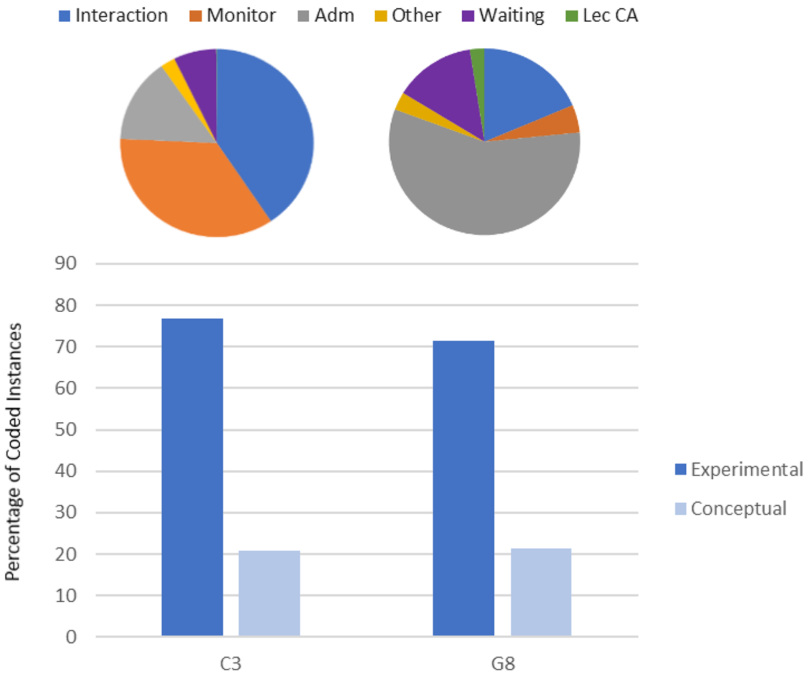


Figure 9. Behavioral maps and the nature of coded instances for two instructors both delivering a REActivities recrystallization lab with 13 students.

interactions between the instructor and students during a lab. This realization suggests that regardless of instructor teaching styles, REActivities ensures a certain degree of fidelity when adopted.

■ CONCLUSIONS

This study analyzed data from seven institutions and eight different instructors teaching first semester undergraduate organic chemistry laboratories to investigate the extent to which the framework of REActivities impacts FOI as it relates to the intended structural and instructional components. The study used the ELITE observational instrument presented for the first time in this paper, designed to measure the FOI of REActivities with the deliberate focus on instructional behaviors during the lab delivery. The ELITE instrument was designed to capture the behavior map for each instructor across institutions that utilized REActivities, revealing a distinct pattern for the pedagogy. This pattern was compared with the behavior maps of both an expert user of the pedagogy (University A, Instructor 1) and with other instructors who delivered the same content either using REActivities or using a traditional approach. FOI was supported through the evaluation of six components deemed critical to the pedagogy. The use of the REActivities workbook, the absence of a prelab lecture, and the student self-start were all confirmed structural components across the participating institutions using the REActivities approach. The instructor behavior maps for institutions utilizing REActivities consistently aligned with the expert user of REActivities and also proved distinguishable from those coded for instructors using traditional lab pedagogies when delivering the same lab content. This suggests that the guided-inquiry materials were robust regardless of teaching style. Upon further analysis of the behavior maps, a more focused look at the nature of the coded student−instructor interactions revealed a consistent degree of discussions involving chemistry concepts when REActivities was employed. Even when an instructor's style had less than expected student−instructor engagement, the same degree of conceptually based discussions ensued in that small amount of time, aligning with instructors who engaged longer with their students. This consistent presence of conceptual discussion across all lab types and each instructor at multiple universities compared strikingly to the near absence of coded conceptual discussion in the expository laboratories during those instructors' student interactions. Thus, our findings support that the intended critical components of the REActivities pedagogy as they relate to the instructor are robustly conserved across various instructors and institutions. This is a promising step toward the claim that REActivities supports opportunities for students to engage in meaningful discussions during a REActivities lab period aligning with inquiry learning practices in a lab setting.

Implications for Teaching

While our previous study looked at the reformed nature of the student−student engagement during a REActivities lab in contrast to expository delivery,[1] this study helps to support the claim that exhaustive instructor training is not necessary to ensure consistent delivery. The data suggests that for those institutions utilizing graduate teaching assistants and large numbers of lab sections with different instructors, REActivities can both alleviate the burden of a prelab lecture while ensuring a comparable delivery experience across lab sections with different instructors. This study also hints that REActivities provides opportunities for students *during* the lab period to think about chemistry concepts, thus creating better opportunities to make connections between underpinnings at the molecular level and the physical activities performed at the bench. One could argue that a prelab lecture or assignment is meant to clarify the lab protocol as it relates to the conceptual underpinnings of the chemistry and thus preclude such discussions during the lab. Although possible, this study shows that such discussions were occurring during the lab at a minimum with the instructor and not delivered to students passively via a prelab when REActivities is utilized; the chemistry concepts are instead incorporated strategically and consistently during the lab period ensuring that the scientific principles are reinforced during the active learning component with the instructor. The overall fidelity of implementation of REActivities supports future studies of the pedagogy when meaningful learning in the lab and student−student discussions are evaluated.

■ AUTHOR INFORMATION

Corresponding Author

Christina Goudreau Collison − *Rochester Institute of Technology, Rochester, New York 14623, United States;* orcid.org/0000-0002-1361-447X; Email: cgcsch@rit.edu

Authors

Dina L. Newman − *Rochester Institute of Technology, Rochester, New York 14623, United States*
Julia Qingli Biehler − *Rochester Institute of Technology, Rochester, New York 14623, United States*
Micaela Nelson − *Rochester Institute of Technology, Rochester, New York 14623, United States*
Paige O'Brien Daly − *Rochester Institute of Technology, Rochester, New York 14623, United States*
Matthew Jackson − *Rochester Institute of Technology, Rochester, New York 14623, United States*
Cameron Isaac − *Rochester Institute of Technology, Rochester, New York 14623, United States*
Madeline Tebrugge − *Rochester Institute of Technology, Rochester, New York 14623, United States*
Jason Anderson − *Monroe Community College, Rochester, New York 14623, United States*
Brian Edelbach − *Monroe Community College, Rochester, New York 14623, United States*
Douglas Tusch − *Rochester Institute of Technology, Rochester, New York 14623, United States*
Jeremy Alan Cody − *Rochester Institute of Technology, Rochester, New York 14623, United States*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jchemed.1c00998

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors would like to thank Gautam Bhattacharyya for his insightful feedback on this project and invaluable advice. This work was supported by the School of Chemistry and Materials Science at Rochester Institute of Technology, the Department of Chemistry at Monroe Community College, and by the National Science Foundation (NSF-IUSE 1625649). The authors are also very grateful to the participating instructors and their institutions.

■ REFERENCES

(1) Collison, C. G.; Kim, T.; Cody, J.; Anderson, J.; Edelbach, B.; Marmor, W.; Kipsang, R.; Ayotte, C.; Saviola, D.; Niziol, J. Transforming the Organic Chemistry Lab Experience: Design, Implementation, and Evaluation of Reformed Experimental Activities (REActivities). *J. Chem. Educ.* **2018**, *95*, 55−61.

(2) Collison, C. G.; Cody, J. A.; Smith, D.; Swartzenberg, J. Formalizing the First Day in an Organic Chemistry Laboratory Using a Studio-Based Approach. *J. Chem. Educ.* **2015**, *92*, 1510−1513.

(3) Collison, C. G.; Cody, J. A.; Stanford, C. An SN1-SN2 Lesson in an Organic Chemistry Lab Using a Studio-Based Approach. *J. Chem. Educ.* **2012**, *89*, 750−754.

(4) Taber, K. S. Revisiting the chemistry triplet: drawing upon the nature of chemical knowledge and the psychology of learning to inform chemistry education. *Chemistry Education Research and Practice* **2013**, *14*, 156−168.

(5) Johnstone, A. H. Macro- and micro-chemistry. *Sch. Sci. Rev.* **1982**, *64*, 377−379.

(6) Johnstone, A. H. Why is science difficult to learn? Things are seldom what they seem. *Journal of Computer Assisted Learning* **1991**, *7* (2), 75−83.

(7) Taber, K. S. Three levels of chemistry educational research. *Chemistry Education Research and Practice* **2013**, *14*, 151−155.

(8) Talanquer, V. A. Macro, submicro, and symbolic: The many faces of the chemistry "triplet". *International Journal of Science Education* **2011**, *33*, 179−195.

(9) Johnstone, A. H. Teaching of chemistry: logical or psychological? *Chemistry Education Research and Practice* **2000**, *1*, 9−15.

(10) Hofstein, A.; Lunetta, V. N. The laboratory in science education: Foundations for the twenty-first century. *Sci. Educ.* **2004**, *88*, 28.

(11) Becker, N.; Rasmussen, C.; Sweeney, G.; Wawro, M.; Towns, M.; Cole, R. Reasoning using particulate nature of matter: An example of a sociochemical norm in a university-level physical chemistry class. *Chemistry Education Research and Practice* **2013**, *14* (1), 81−94.

(12) Stanford, C.; Moon, A.; Towns, M.; Cole, R. Analysis of instructor facilitation strategies and their influences on student argumentation: A case study of a POGIL physical chemistry classroom. *J. Chem. Educ.* **2016**, *93* (9), 1501−1513.

(13) Yackel, E.; Cobb, P. Sociomathematical Norms, Argumentation, and Autonomy in Mathematics. *J. Res. Math. Educ.* **1996**, *27* (4), 458−477.

(14) Becker, N.; Stanford, C.; Towns, M.; Cole, R. Translating across macroscopic, sub-microscopic, and symbolic levels: The role of instructor scaffolding in an inquiry-oriented physical chemistry class. *Chemistry Education Research and Practice* **2015**, *16*, 769−785.

(15) Stanford, C.; Moon, A.; Towns, M.; Cole, R. The Impact of Guided Inquiry Materials on Student Representational Level Understanding of Thermodynamics. In *Engaging Students in Physical Chemistry*; American Chemical Society: Washington, D.C., 2018; pp 141−168.

(16) Winn, W. Instructional Design and Situated Learning: Paradox or Partnership? *Educ. Technol.* **1993**, *33* (3), 16−21.

(17) Hunnicutt, S. S.; Grushow, A.; Whitnell, R. Guided-Inquiry Experiments for Physical Chemistry: The POGIL-PCL Model. *J. Chem. Educ.* **2015**, *92*, 262−268.

(18) Carmel, J. H.; Herrington, D. G.; Posey, L. A.; Ward, J. S.; Pollock, A. M.; Cooper, M. M. Helping Students to "Do Science": Characterizing Scientific Practices in General Chemistry Laboratory Curricula. *J. Chem. Educ.* **2019**, *96*, 423−434.

(19) Clark, T. M.; Ricciardo, R.; Weaver, T. Transitioning from Expository Laboratory Experiments to Course-Based Undergraduate Research in General Chemistry. *J. Chem. Educ.* **2016**, *93*, 56−63.

(20) Lave, J.; Wenger, E. *Situated Learning: Legitimate Peripheral Participation*; Cambridge University Press: New York, 1991.

(21) Hung, D. Situated Cognition and Problem Based Learning: Implications for Learning and Instruction with Technology. *J. Interact. Learn. Res.* **2002**, *13* (4), 393−414.

(22) Brown, J. S.; Collins, A.; Duguid, P. Situated cognition and the culture of learning. *Educational researcher* **1989**, *18* (1), 32−42.

(23) Leach, J.; Scott, P. Individual and Sociocultural Views of Learning in Science Education. *Science & education* **2003**, *12* (1), 91−113.

(24) Herron, J. D.; Nurrenbern, S. C. Chemical Education Research: Improving Chemistry Learning. *J. Chem. Educ.* **1999**, *76*, 1353−1361.

(25) Jalil, P. A. A Procedural Problem in Laboratory Teaching: Experiment and Explain, or Vice-Versa? *J. Chem. Educ.* **2006**, *83* (1), 159−163.

(26) Landis, C. R.; Peace, G. E.; Scharberg, M. A.; Branz, S.; Spencer, J. N.; Ricci, R. W.; Zumdahl, S. A.; Shaw, D. The New Traditions Consortium: Shifting from a Faculty-Centered Paradigm to a Student-Centered Paradigm. *J. Chem. Educ.* **1998**, *75* (6), 741−744.

(27) Haight, G. P. Bringing Undergraduates to the Chemical Frontier. *J. Chem. Educ.* **1967**, *44* (12), 766−767.

(28) Ricci, R. W.; Ditzler, M. A. Discovery Chemistry: A Laboratory-Centered Approach to Teaching General Chemistry. *J. Chem. Educ.* **1991**, *68* (3), 228−231.

(29) Ricci, R. W.; Dizler, M. A.; Jarret, R.; McMaster, P.; Herrick, R. The Holy Cross Discovery Chemistry Program. *J. Chem. Educ.* **1994**, *71* (5), 404−405.

(30) Buck, L. B.; Bretz, S. L.; Towns, M. H. Characterizing the Level of Inquiry in the Undergraduate Laboratory. *J. Coll. Sci. Teach.* **2008**, *38* (1), 52−58.

(31) Furtak, E. M. The Problem with Answers: An exploration of guided scientific inquiry teaching. *Sci. Educ.* **2006**, *90* (3), 453−467.

(32) Berry, A.; Gunstone, R.; Loughran, J.; Mulhall, P. Using Laboratory Work for Purposeful Learning About the Practice of Science. *Research in Science Education - Past, Present, and Future* **2002**, 313−318.

(33) Xu, H.; Talanquer, V. Effect of the Level of Inquiry on Student Interactions in Chemistry Laboratories. *J. Chem. Educ.* **2013**, *90* (1), 29−36.

(34) Schoffstall, A. M.; Gaddis, B. A. Incorporating Guided-Inquiry Learning into the Organic Chemistry Laboratory. *J. Chem. Educ.* **2007**, *84* (5), 848.

(35) Stains, M.; Vickrey, T. Fidelity of Implementation: An overlooked yet critical construct to establish effectiveness of evidence-based instructional practices. *CBE Life Sci. Educ.* **2017**, *16* (1), rm1.

(36) Smith, M. K.; Jones, F. H.; Gilbert, S. L.; Wieman, C. E. The Classroom Observation Protocol for Undergraduate STEM (COPUS): a new instrument to characterize university STEM classroom practices. *CBE Life Sci. Educ.* **2013**, *12* (4), 618−627.

(37) Velasco, J. B.; Knedeisen, A.; Xue, D.; Vickrey, T. L.; Abebe, M.; Stains, M. Characterizing Instructional Practices in the Laboratory: The laboratory observation protocol for undergraduate STEM. *J. Chem. Educ.* **2016**, *93* (7), 1191−1203.

(38) Missett, T. C.; Foster, L. H. Searching for evidence-based practice: a survey of empirical studies on curricular interventions measuring and reporting fidelity of implementation published during 2004−2013. *J. Adv. Acad.* **2015**, *26*, 96−111.

(39) Henderson, C.; Dancy, M. H. Impact of physics education research on the teaching of introductory quantitative physics in the United States. *Phys. Rev. Spec Top Phys. Educ Res.* **2009**, *5*, 020107.

(40) Turpen, C.; Finkelstein, N. D. Not all interactive engagement is the same: variations in physics professors' implementation of peer instruction. *Phys. Rev. Phys. Spec Top Educ Res.* **2009**, *5*, 020101.

(41) Chase, A.; Pakhira, D.; Stains, M. Implementing process-oriented, guided-inquiry learning for the first time: adaptations and short-term impacts on students' attitude and performance. *J. Chem. Educ.* **2013**, *90*, 409−416.

(42) Daubenmire, P. L.; Bunce, D. M.; Draus, C.; Frazier, M.; Gessell, A.; Van Opstal, M. T. During POGIL implementation the professor still makes a difference. *J. Coll Sci. Teach.* **2015**, *44*, 72−81.

(43) Mowbray, C. T.; Holter, M. C.; Teague, G. B.; Bybee, D. Fidelity criteria: development, measurement, and validation. *Am. J. Eval.* **2003**, *24*, 315−340.

(44) Century, J.; Rudnick, M.; Freeman, C. A framework for measuring fidelity of implementation: a foundation for shared language and accumulation of knowledge. *Am. J. Eval.* **2010**, *31*, 199−218.

(45) Carroll, C.; Patterson, M.; Wood, S.; Booth, A.; Rick, J.; Balain, S. A conceptual framework for implementation fidelity. *Implement. Sci.* **2007**, *2*, 40.

(46) O'Donnell, C. L. Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Rev. Educ Res.* **2008**, *78*, 33−84.

(47) West, E. A.; Paul, C. A.; Webb, D.; Potter, W. H. Variation of instructor-student interactions in an introductory interactive physics course. *Phys. Rev. Spec. Top. Phys. Educ. Res.* **2013**, *9*, 010109.

(48) Generalized Observation and Reflection Platform. https://gorp.ucdavis.edu/ (accessed August 2021).

(49) Nelson, M. C.; Cordray, D. S.; Hulleman, C. S.; Darrow, C. L.; Sommer, E. C. A procedure for assessing intervention fidelity in experiments testing educational and behavioral interventions. *J. Behav Health Serv Res.* **2012**, *39*, 374−396.

(50) Institute of Education Sciences. *Common Guidelines for Education Research and Development: A Report from the Institute of Education Sciences and the National Science Foundation*; U.S. Department of Education: Washington, D.C.; 2013.

(51) Hedges, L. V. Distribution theory for Glass' estimator of effect size and related estimators. *J. Educ. Stat.* **1981**, *6*, 107−128.

(52) Hedges, L. V.; Olkin, I. *Statistical Methods for Meta-Analysis*; Academic Press, Inc.: Orlando, FL, 1985.