ARTICLE

Backward Design As a Framework for Strengthening CURE Development

Kelly Y. Neiles, *St. Mary's College of Maryland*
Rebecca A. Hunter, *College of New Jersey*
Maury E. Howard, *Virginia Wesleyan University*
Daniel F. Scott, *Centre College*
Kimberley A. Frederick, *Skidmore College*

Abstract

Course-based undergraduate research experiences (CUREs) increase student engagement, persistence in STEM, and development of scientific practices. Although recent publications address implementation logistics, few studies specify curriculum development methods that result in inclusion of scientific practices and authentic research while maintaining necessary scaffolding. This article describes a professional development experience using backward design to support such curriculum development. Four analytical co-principal investigators engaged in backward design modules leading to the development of the μCURE project. Evaluation of the μCURE included student perceptions and analysis of teaching materials using modified inquiry and Three-Dimensional Learning Assessment Protocol (3D-LAP) rubrics. Results showed high student perceptions of learner-centeredness, nearly authentic levels of inquiry, and increase in the presence of scientific practice prompting. The backward design method was found to strengthen high-quality CURE development.

Keywords: *authentic assessment; chemistry; course-based undergraduate research (CURE); faculty development*

doi: 10.18833/spur/9/3/6

Engaging students in authentic research experiences is a proven, high-impact practice useful for training students to think and act as scientists (Eagan et al. 2013; Lopatto 2007; Russell, Hancock, and McCullough 2007; Sadler and McKinney 2010; Weaver, Russell, and Wink 2008). These experiences can enhance students' understanding of the scientific method and provide vital opportunities for developing foundational scientific practices such as formulating explanations from evidence and planning and carrying out investigations (Carter, Mandell, and Maton 2009; Chang 2002; Russell et al. 2007). As opposed to the traditional and often resource-heavy apprenticeship model, course-based undergraduate research experiences (CUREs) serve as a scalable and equitable pedagogical solution, effectively broadening access to authentic research for a greater proportion of students by grounding the experiences in courses in which all enrolled students can benefit. CUREs offer the capacity to involve many students in research (e.g., Rowland et al. 2012) and can be integrated into introductory-level courses (Dabney-Smith 2009; Harrison et al. 2011). Therefore, they have the potential to introduce scientific practices to a wider swath of students earlier in their academic careers, reducing entry barriers to research (Hunter, Laursen, and Seymour 2007; Provost 2022).

Recent consensus on the features that define a CURE describes the integration of five key dimensions that, when implemented collectively, distinguish it from traditional laboratory instruction and other inquiry-based experiences. These are: (1) the element of discovery, in which the outcome is unknown to both students and instructors; (2) the contribution of student work to broadly relevant questions important to a stakeholder community; (3) the emphasis on collaboration; (4) the requirement that students engage

in scientific practices; and (5) the necessary inclusion of iteration, which mandates time for students to fail, troubleshoot, and revise experiments to mirror the authentic scientific process (Corwin et al. 2014; Corwin et al. 2018; Dolan and Weaver 2021; Provost 2022; Provost, Bell, and Bell 2019). Ensuring that the first three of these dimensions are present results from the instructor identifying an appropriate topic area for the CURE and planning logistics that facilitate collaboration. The last two dimensions come from careful curriculum design and implementation.

Close to a decade ago, the CURE community called for a shift toward theoretical framework–driven design and assessment of CUREs (Cooper, Soneral, and Brownell 2017; Corwin et al. 2014). Indeed, in 2016, Dolan reported that few studies existed on the design and implementation of CURES, or on how instructors learned to teach them. Since then, several publications have addressed this gap (Dolan and Weaver 2021; Provost 2022; Provost, Bell, and Bell 2019), although they have focused predominantly on logistical mechanics, such as overcoming implementation barriers related to cost, time, and scaling access, some of the most prevalent barriers to the implementation of CUREs. Unfortunately, this has left a critical gap. Few studies have presented theoretical frameworks or specific curriculum development methods needed to effectively design CUREs to include scientific practices and intentional learning experiences that provide students with the opportunity to engage in an iterative process. This is concerning, because studies have shown that when instructors design experiences requiring higher levels of inquiry, such as CUREs, they often inadvertently remove the scaffolding that prompts students to engage in scientific practices, the exact thing the CUREs are meant to teach (Van Wyk et al. 2025). The implication of these studies is that rather than just assuming the higher level of inquiry equates to increased engagement with scientific practices, instructors must intentionally build opportunities to use scientific practices while also maintaining the element of discovery needed for the experience to be an authentic research experience. This is a fine line to walk, one that is often difficult without the use of specific curriculum design methods.

One curriculum design framework that has been proposed to address barriers in chemistry laboratories (Neiles and Arnett 2021), and in CUREs specifically (Cooper et al. 2017), is a method called backward design (Wiggins and McTighe 2011), in which the curriculum developer proceeds through a three-step process that, when employed, facilitates a more intentionally designed and scaffolded curriculum (Neiles and Arnett 2021). Through this process the instructor identifies the desired learning outcomes and ensures that those components are present in the assessments and learning experiences included in the CURE. Additionally, the instructor can use the process to self-check whether their pedagogical decisions are student-centered and transparent, creating a learning environment that promotes success for all students.

The central goal of this article is to describe a professional development (PD) experience built around the principles of backward design, with the intent of guiding faculty in creating CUREs grounded in scientific practices and student-centered pedagogy. The authors ask: Can a backward design process lead to CUREs that are designed in a way that prompts for engagement in scientific practices and authentic research while maintaining the scaffolding necessary to allow students to meet those goals? The authors hypothesize that the combination of curriculum design theory, collaborative development, and practical design tools will empower future faculty to create CUREs with these desired elements.

Development Process

Overview

The backward design PD experience was developed as part of a National Science Foundation Division of Undergraduate Education Improving Undergraduate STEM Education (DUE-IUSE) grant (award 2215768). In this grant four analytical co-principal investigators were led by a chemical education (ChemEd) co-principal investigator through the backward design process to develop a networked CURE focused on microfluidic paper-based analytical devices (the μCURE project). To develop the μCURE, the group met monthly over the course of one semester and one summer to collaboratively engage in the backward design process through PD modules. After these initial modules, the analytical principal investigators engaged in an iterative collaboration process that consisted of yearly μCURE implementation at their home institutions, followed by summer meetings to debrief the implementation, again discuss backward design elements, and develop a work plan for μCURE refinement. The resulting μCURE project focused on the process of translating a traditional solution-phase spectrophotometric assay into a microfluidic paper analytical device.

The general flow of the developed CURE took place in five stages. Conceptual planning for this research occurred outside of class, beginning with students conducting a literature search to identify a viable research topic and subsequently developing a formal research proposal, which instructors evaluated for scientific soundness and executability. Once the project's laboratory work began in the final third of the semester, the scientific process involved students depositing reagents onto filter paper disks (or wax printed devices) and drying them. Standard or sample solutions were then added to the paper, initiating a color-producing reaction, the intensity of which was proportional to the analyte concentration. Students captured images of the results with a smartphone and analyzed them with digital imaging software to measure RGB values,

FIGURE 1. QR Code for Supplemental Information

Note: Also available at https://tinyurl.com/4d8rewcv and as a Google document upon request to the authors.

which they then used to produce calibration curves. This experimental investigation was highly iterative, reinforced by structured weekly progress reports in which students reflected on their results and proposed adjustments to their experimental plan. Finally, using the gathered data, students determined figures of merit (e.g., detection limit, sensitivity), compared these metrics against the needs of their sample analysis to assess goal achievement, and communicated their findings during a joint virtual poster session. The full description of the resulting μCURE project can be found in a manuscript published by Frederick et al. (2025).

Backward Design

The backward design PD modules were completed via a Google website during spring and summer 2023. Four backward design modules, developed based on the work of Neiles and Arnett (2021), were included in the PD:

Module 1 Backward design Step 1: Identification of learning outcomes
Module 2 Backward design Step 2: Evidence of student learning (assessment)
Backward design Step 3: Creating the teaching plan
Module 3 Scaffolding of learning experiences
Module 4 Developing a timeline and necessary teaching materials

A fifth module on the use of transparent teaching principles also was included, but will not be discussed in the present article as it was not the focus of this work. Each of the four modules began with a detailed description of the relevant step in backward design and then progressed to an assignment done via Google documents. The full set of modules and assignments can be accessed with the QR code or URL in Figure 1.

The analytical principal investigators took two weeks to work through each module. Then, the ChemEd co-principal investigator reviewed their work and provided feedback in the following week. This feedback would generally focus on trends found in the work. For example, after the analytical principal investigators created their first set of learning outcomes, the ChemEd co-principal investigator highlighted similarities and differences in the learning outcomes produced. This led to the principal investigators agreeing upon the initial set of learning outcomes to be implemented in the μCURE and creating a common understanding of what those learning outcomes would entail. Finally, the group would meet to discuss the module and make any necessary collaborative decisions for the μCURE to be networked across institutions. The following sections briefly describe the four modules and corresponding assignments.

Module 1: Identification of Learning Outcomes. Module 1 centered on the systematic articulation of learning outcomes for the μCURE (supplemental information, sections 2 and 3). The assignment for this module guided the analytical co-principal investigators through the process of documenting the course and student context (which would influence future pedagogical decisions), including how to identify factors such as prerequisite knowledge. Instructors were directed to initially brainstorm broad goals for the laboratory experience, including any desired scientific practices. The analytical co-principal investigators used the Next Generation Science Standard (NGSS) list of nine scientific practices (National Research Council 2012) to identify six practices for the μCURE.

- NGSS Scientific practice (SP) 3: Planning and carrying out investigations
- NGSS SP 4: Analyzing and interpreting data
- NGSS SP 5: Using mathematical and computational thinking
- NGSS SP 6: Constructing explanations and engaging in arguments from evidence
- NGSS SP 7: Evaluating information
- NGSS SP 8: Defining problems and designing solutions

The three NGSS SPs not selected were SP1: Asking questions; SP2: Developing and using models; and SP9: Communicating information. Although the co-principal investigators believed these were important scientific practices, it was determined that they would not be the focus of this particular μCURE project. The six selected scientific practices, along with other initial goals, were refined by defining the anticipated level of student proficiency (categorized as foundational, developing, or mastery) before consolidating them into a concise set of formal outcomes. The module prescribed a specific format for writing these formal outcomes ("students will + observable verb + conditions of the learning") to ensure they were measurable, deliberately excluding nonobservable terms such as "understand" or "know." Finally, the conditions of learning were integrated to provide necessary context and specificity, coupling the desired scientific behavior (the verb) with relevant content or techniques. Below are the

five learning outcomes identified through Module 1 (and refined by future iterations). To assist the reader in identifying the learning outcome elements, the observable verbs are bold, and the conditions of learning are italicized.

Upon completion of the μCURE experience students would be able to:

- **Identify** and **assess** *relevant [to the μCURE project]** literature sources.
- **Propose** a viable experimental plan *to answer a well-defined scientific question based on literature information and experimental results [relevant to the μCURE]*.
- **Apply** appropriate methods of data analysis *to interpret [μCURE] experimental results*.
- **Evaluate** multiple pieces of *[μCURE] experimental data* to support conclusions.
- **Contribute** to a team by working collaboratively toward common *[μCURE]* goals.

*Square brackets indicate an assumed tie to the μCURE project within a learning objective. The assumption holds as these learning outcomes were provided to students within the μCURE project context.

It is important to note that just because a scientific practice is desired, it does not necessarily need to be its own separate learning outcome. Instead, the scientific practices often show up in the later backward design steps as assessments or learning activities. The outcomes written by the analytical co-principal investigators exemplify learning outcomes written through the backward design approach because they are clear, concise, and include an observable verb and conditions of learning. A well written outcome is vitally important because it acts as the foundation upon which the remaining modules are built.

Module 2: Evidence of Student Learning (Assessment). Module 2 guided the analytical co-principal investigators through the second step in the backward design framework, focusing on how instructors determined if students had achieved the desired learning outcomes through the identification of student learning evidence (SI sections 4 and 5). Upon completing Module 2, instructors were expected to have identified appropriate learning artifacts (evidence), considered the criteria by which they would evaluate that evidence, and selected and/or developed corresponding assessment methods (often rubrics). To do this, the analytical co-principal investigators first returned to the observable verb for each learning outcome to ensure that the evidence they had selected during Module 2 reflected the action intended by the outcome. For example, in the outcome, "Evaluate multiple pieces of [μCURE] experimental data to support conclusions," the observable verb "evaluate" implied that students would need to make a judgment about the methods they selected based on specific criteria and standards. The students should encounter an opportunity to provide evidence of these skills, and be assessed using a metric that looks for the relevant criteria. In the μCURE project, students provided evidence for this outcome in their final poster session and were assessed with an Enhancing Learning by Improving Process Skills in STEM (ELIPSS) critical thinking rubric (ELIPSS 2025; Reynders et al. 2019). The criteria on this rubric were relevant to the learning outcome because they included the following:

- Identifying the goal: Determined the purpose/context of the argument or conclusion that needed to be made.
- Evaluating: Determined the relevance and reliability of information that might be used to support the conclusion or argument.
- Analyzing: Interpreted information to determine meaning and to extract relevant evidence.
- Synthesizing: Connected or integrated information to support an argument or reach a conclusion.

When used effectively, these criteria and the ratings related to them provided a score relevant to the learning outcome.

Module 3: Scaffolding of Learning Experiences. Module 3 acted as a crucial pre-step within the third phase of the backward design framework, focusing on the essential curricular development needed to support student mastery of learning outcomes (SI sections 6 and 7). This module guided the analytical co-principal investigators through the process of establishing a clear understanding of student proficiency by documenting the expected knowledge students possessed both before and at the commencement of the μCURE project. The objective was to ensure that learning activities throughout the μCURE (and even prior to the μCURE) provided appropriate scaffolding, building student proficiency through a sequence of iterative experiences that progressed from smaller foundational units to a larger, complex whole. Upon completing this module, instructors had articulated the specific "touchpoints" that guided students effectively from their starting point to the desired final outcome. Again, using the example outcome, "Evaluate multiple pieces of [μCURE] experimental data to support conclusions," the analytical co-principal investigators determined to what degree their specific student population had prior knowledge of this outcome. In most cases, the students' prior knowledge before entering the course was limited to their introductory chemistry courses, in which they had some basic experience with drawing conclusions from data by using data to inform the next steps of the experiment. In previous labs within the course, students received instruction on relevant analytical topics such as figures of merit, statistics, and validation. This meant that scaffolding for this outcome within the μCURE could assume a basic understanding of drawing conclusions from data and how to do that using the newly acquired analytical knowledge and skills.

TABLE 1. Sample Instructional Sequence Identified for Learning Outcome 4

Touchpoint	Teaching material
1. (Candy lab) Students receive instruction on interpreting statistical results in class.	In-lab activity
2. (Candy lab) Students draw conclusions based on hypothesis testing in the candy lab.	Assignment
3. (Iron lab) Students receive direct instruction on how to use data to plan next experiment in iron lab.	Slides/activity
4. (Iron lab) Students practice interpreting their own data and making an experimental plan for week 2 of iron lab.	Assignment rubric (ELIPSS critical thinking)
5. (μCURE lab) Starting in week 2 of μCURE, each week students evaluate and interpret their data, draw conclusions, and make a plan based on those conclusions.	Assignment
6. (μCURE lab) Students receive feedback on the accuracy of data interpretation and appropriateness of experimental plan.	Rubric (ELIPSS critical thinking)

Module 4: Developing a Timeline and Necessary Teaching Materials. Module 4 guided the analytical co-principal investigators in completing the third step of backward design, leveraging the outcomes (Module 1), evidence (Module 2), and scaffolding plan (Module 3) to construct the instructional sequence (SI sections 8 and 9). The primary objective of this module was for instructors to identify the specific teaching materials and learning activities necessary to allow students to achieve the desired outcomes and successfully show the evidence of learning. The analytical co-principal investigators were tasked with developing a detailed timeline for the μCURE experience, integrating the scaffolded touchpoints identified in Module 3 to ensure the curriculum progressed in a logical, stepwise manner to meet all identified learning outcomes. Returning to the example outcome, "Evaluate multiple pieces of [μCURE] experimental data to support conclusions," the analytical co-principal investigators created scaffolded touchpoints contextualized for their own institution. Table 1 shows one institution's example of those touchpoints along with the relevant teaching materials.

This instructor's example illustrates the scaffolding that happens to prepare for meeting the learning outcome both before the μCURE lab (1 through 4), and within the lab (5 and 6), highlighting the need for careful pedagogical planning across the semester.

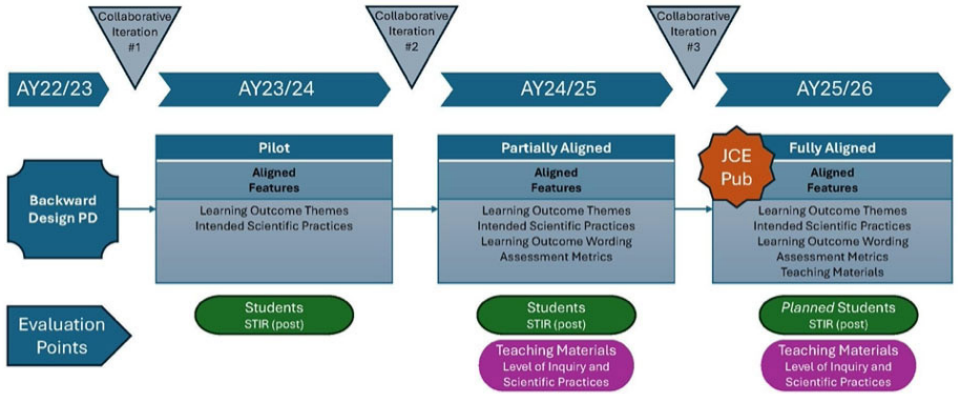
Collaborative Iteration

The full timeline for development of the final μCURE as described in Frederick et. al (2025) was from fall 2023 through summer 2025 and included three phases: (1) pilot, (2) partially aligned, and (3) fully aligned. The phases, alignment of μCURE characteristics between institutions, and evaluation of the μCURE at each phase are visualized in Figure 2. Although total alignment between institutions was not a goal, collaborative iteration discussions held after each implementation resulted in further refinement of the experience and ultimately alignment across institutions. In August of 2025, the fully aligned μCURE was published in the *Journal of Chemical Education* (Frederick et al. 2025).

Collaborative Iteration 1. Through the four modules described, the analytical co-principal investigators created μCURE experiences contextualized for their home institutions. After development, the ChemEd co-principal investigator reviewed the materials and helped the analytical co-principal investigators identify a common set of learning outcome themes. Through discussion the group ensured that these common themes encompassed the scientific practices agreed upon by the co-principal investigators during Module 1. They then came up with a plan to finalize development of the μCURE teaching materials prior to implementation at each of their institutions. Therefore the pilot phase μCUREs were aligned in learning outcome themes and the intended scientific outcomes, but were not necessarily aligned in wording of outcomes, types of assessments utilized, types of scaffolding, and types of teaching materials present in the μCURE. The pilot μCUREs were implemented during the 2023–24 academic year. As each analytical co-principal investigator completed the teaching materials (prior to implementation), those materials were shared with the ChemEd co-principal investigator for cursory evaluation. Full evaluation of the materials came after the pilot and is described in the "Evaluation" section.

*Collaborative Iteration 2***.** In the summer of 2024 the co-principal investigators met to debrief the pilot and return

FIGURE 2. Full Timeline for Development of Published μCUR



Note: STIR, Science Teacher Inquiry Rubric.

to their backward design foundation. During this meeting the ChemEd co-principal investigator provided cursory feedback on whether each of their μCURE projects had scaffolding, teaching materials, and assessments that aligned with the previously articulated learning outcomes. A number of disconnects were identified and discussed. For example, students were receiving effective scaffolding (visible in the touchpoints included in the μCURE) on how to "analyze and interpret data"; however, "arguing conclusions from data" was mostly missing from the teaching materials. Based on these discussions, and others regarding the mechanics and logistics of the μCURE, the co-principal investigators created a work plan for the summer to reformat and develop necessary components of the μCURE. They determined it would be best for the projects if they aligned their learning outcomes and use of specific assessments across institutions. The resulting μCUREs were implemented at their home institutions. Chosen assessments and teaching materials were shared with the ChemEd co-principal investigator for μCURE evaluation.

Collaborative Iteration 3. The final collaborative iteration meeting happened in the summer of 2025 when the co-principal investigators met after the implementation of their partially aligned μCUREs to again debrief their implementation. It was at that meeting, and in preparation for final publication of the μCURE project, that the analytical co-principal investigators determined it would be best if they aligned their teaching materials across institutions. Additionally, they realigned their assessment methods to better meet the needs of the fully aligned μCURE. At this point, the μCURE project was submitted and accepted for publication. It was published in August of that summer, although the co-principal investigators continue to refine and improve the project and plan to implement yet another iteration during the 2025–26 academic year.

Evaluation

To answer the research question, "Can a backward design process lead to CUREs that are designed in a way that prompts for engagement in the scientific practices and authentic research while maintaining the scaffolding necessary to allow students to meet those goals?" the μCUREs were evaluated by two forms of data: student perceptions and μCURE teaching materials.

Student Perceptions

During the pilot and partially aligned implementations, the μCUREs were characterized using the Science Teacher Inquiry Rubric (STIR; Bodzin and Beerer 2003). The STIR rubric was selected because it was developed and validated by Bodzin and Beerer based upon the National Science Education Standards (NSES) essential features of inquiry instruction. STIRs were completed by students after the μCURE project was completed at each implementation. The students were asked to rate the extent to which they had an opportunity during the μCURE to: (1) engage with a scientific question; (2) plan investigations to gather evidence in response to questions; (3) formulate conclusions and/or explanations from evidence to address scientific questions; and (4) communicate and justify their proposed conclusions and/or explanations. Responses ranged from a scale of 1 (learner centered) to 4 (teacher centered), or "no evidence observed." During the pilot implementation two institutions were able to collect these data from students (n = 27) and during the partially aligned implementation three institutions were able to collect these data (n = 34). The data were aggregated and used to characterize laboratory instruction on a continuum from teacher- to learner-centered as perceived by the students.

Teaching Materials

Teaching materials for the μCUREs were evaluated at two

TABLE 2. Summary of Teaching Materials Evaluated for Partially Aligned μCUREs

Institution 1	Institution 2	Institution 3	Institution 4
Micro-CURE guidelines	Individual literature search activity	Micro-CURE device design	Micro-CURE device design
Week 1 pre-lab	Group literature search activity	Library exercise	Make your own microfluidics device pt 1
Annotated bibliography assignment	Annotated bibliography assignment	Annotated bibliography assignment	Annotated bibliography assignment
Weekly progress reports	Weekly progress reports	Weekly progress reports	Weekly progress reports
Group contract	Proposal draft	Research team contract	Research group contract
Moles and concentration	Proposal final		

TABLE 3. Summary of Teaching Materials Evaluated for Fully Aligned μCURE

	Teaching materials
Preassignments	Microfluidics, moles, and concentration
	Adapting a procedure from the literature
	Analytical method validation and arguing conclusions from data
μCURE assignments	Project description and timeline
	Literature search activity
	Group contract
	Group literature search
	Project proposal draft
	Final proposal
	Weekly progress reports
	Final report

points: after they were partially aligned (following Collaborative Iteration 2) and after they were fully aligned (following Collaborative Iteration 3). Given that the partially aligned materials were not similar across institutions, they were evaluated by institution. Table 2 provides a summary of what materials were evaluated for the μCURE at each institution. Each assignment listed could have multiple "artifacts" used in the evaluation (handout, presentation slides, rubrics, etc.). In a few cases, the same teaching materials were used at multiple institutions, for example, the annotated bibliography assignment.

The fully aligned μCURE had a common set of teaching materials used by all instructors and an additional set of preassignments, which could be used by instructors to scaffold skills leading up to the μCURE. Table 3 provides a summary of what materials were evaluated for the fully aligned and published μCURE.

The partially and fully aligned teaching materials were analyzed to determine their level of inquiry and the extent to which they provided opportunities for students to engage in scientific practices. The evaluation of inquiry level was conducted using a modified version of the inquiry rubric designed by Bruck, Bretz, and Towns, which defines levels of inquiry by the amount of structure provided to students across six core laboratory characteristics: problem/question, theory/background, procedure/design, results analysis, results communication, and conclusions (Bruck, Bretz, and Towns 2008; Van Wyk et al. 2024). The evaluation involved assigning one of three subcodes—"provided," "partially provided," or "not provided"—to

each of the six characteristics based on the content present in the μCURE projects. Ultimately, the overall level of inquiry for the projects at each phase of development (partially aligned and fully aligned) was determined by matching the resulting pattern of these subcodes across all six characteristics to the corresponding predefined inquiry levels in the modified rubric, thereby assessing the degree to which the materials facilitated student-directed decisions versus traditional, faculty-driven laboratory instruction methods.

The evaluation of laboratory teaching materials for the presence of opportunity for students to engage with scientific practices was conducted with a version of the Three-Dimensional Learning Assessment Protocol (3D-LAP), which was specifically modified to characterize the prompted opportunities for students to engage in the science and engineering practices outlined in the Framework for K–12 Science Education (Laverty et al. 2016; National Research Council 2012; Van Wyk et al. 2025). In this modification, specific criteria were articulated for each scientific practice. For example, for the scientific practice "constructing explanations and engaging in arguments from evidence," four criteria were articulated. The presence of criteria for each of the nine scientific practices was evaluated for each of the teaching materials summarized in Tables 2 and 3. Based on the criteria identified, the materials were categorized at one of five levels: (1) absent: all criteria for the scientific practice were absent; (2) base criteria: only the first criterion for the prompting of a science practice was met; (3) partial task only: only two of the four criteria defining the science practice were met; (4) task only: all criteria were met except the final criterion, which typically requires a student to provide a justification or explanation for their work; or (5) complete prompting: all criteria necessary for the explicit prompting of a science practice were met in the laboratory materials. This process resulted in identification of the extent to which explicit opportunities were provided for students to engage in a given practice.

Data Analysis

The coding and interrater reliability process described here follows that of the work in which the coding was originally developed (Gwet 2002; Van Wyk et al. 2024; Van Wyk et al. 2025). A subset, 33 percent, of the teaching materials from both the partially and fully aligned μCUREs were coded for level of inquiry and presence of scientific practices by the methods described previously and using the codebooks shared in the original work (Van Wyk et al. 2024; Van Wyk et al. 2025). Negotiated agreement between the ChemEd co-principal investigator and two research assistants was used at this stage to discuss differences in coding, and discrepancies were clarified based on conversations. Once coding was consistent, all teaching materials listed in Tables 2 and 3 were coded by one of the three coders. Interrater reliability measures were calculated between all possible pairings of the coders using Gwet's AC1 for all of the teaching materials coded by all three coders (n = 12 out of 36 total teaching materials). Interrater reliability was established with all pairing at a level of $\kappa = 0.80$ or higher, indicating strong reliability (McHugh 2012).

Results and Discussion

Student Perceptions

The STIR rubric results for the pilot and partially aligned implementation are presented in Figure 3.

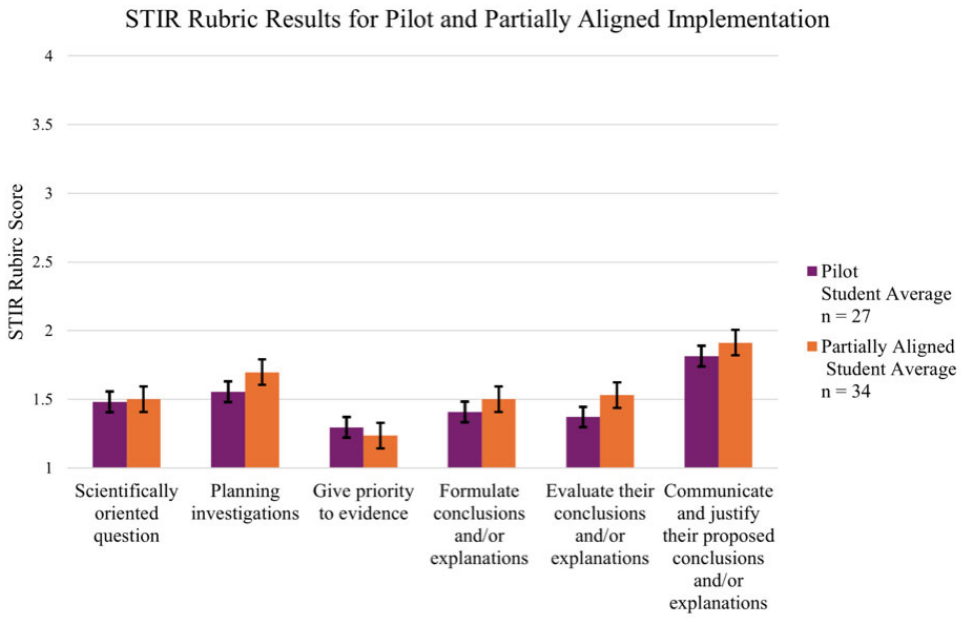
Across both implementation years, students perceived the overall μCURE to be learner-centered, as demonstrated by all average scores remaining well below the midpoint of the scale. A two-tailed independent samples t test confirmed that there were no significant differences in student answers between student responses during the pilot and during the partially aligned implementation on any rubric category ($p > 0.05$), suggesting stable student perceptions across the implementation phases. Although all averages trended toward learner-centeredness, the highest scores were consistently observed in the category "Communicate and justify their proposed conclusions and/or explanations," a finding that is possibly attributable to the required structure of creating a poster. Similarly, the "Planning investigations" category scored slightly higher than others, potentially reflecting guiding questions posed to students during their proposal development. Overall, the aggregated STIR data confirm that students perceived the μCURE approach to be highly learner-centered, with all average scores remaining consistently on the student-centered side of the scale. The stability of these perceptions across the two implementations suggests that the μCURE structure consistently supported perceived learner-centered engagement in scientific practices.

Inquiry Levels

The inquiry rubric findings on the partially and fully aligned teaching materials are presented in Table 4.

The evaluation with a modified inquiry rubric confirmed that the μCURE materials successfully fostered experiences nearing the goal level of authentic inquiry. Specifically, the overall level of inquiry for the fully aligned μCURE was determined to be 2.5, indicating an experience "nearing authentic inquiry," which aligns with the typical pedagogical goals for CUREs. Similarly, during the partially aligned phase, three out of four institutions were rated at level 2.5, with one reaching level 1.5 ("nearing open inquiry"), demonstrating that the majority of implementations were moving toward authentic inquiry. This authentic context was created because, although the broad topic area—translating a traditional spectrophotometric assay into a paper microfluidic format—was provided, the

FIGURE 3. STIR Results



Note: STIR, Science Teacher Inquiry Rubric.

specific analyte selection was left open, resulting in the problem/question characteristic being coded as "partially provided." These results indicate that the fully aligned μCURE placed sufficient decision-making autonomy on the student across the five rubric areas. These findings confirm that the backward design process, along with collaborative iteration, successfully produced CURE materials that aligned with the target level of inquiry.

Scientific Practices

As has been noted, the teaching materials were evaluated for the presence of six of the nine NGSS scientific practices (SPs). Analysis was actually completed for all nine SPs to get a full picture of which NGSS SPs were present. Results broken down by each teaching material and all nine scientific practices can be found in the supplemental materials (SI Section 10; see Figure 1). Table 5 includes the overall presence of scientific practices for each institution's partially aligned μCURE and the fully aligned μCURE.

Across the partially aligned μCUREs, the materials were found to have evidence of the six intended practices (SP3 through SP8). However, a key observation was that many of these partially aligned μCUREs failed to prompt the final critical criteria, which asks students for justification (categorized as "task only"). The subsequent iterative collaboration resulted in the fully aligned μCURE, for which the opportunities for explicit prompting on the final criteria were increased; five out of the six intended scientific practices (SP3, SP4, SP5, SP6, and SP7) achieved "complete prompting," with only one practice (SP8) coded as "task only." This progression to nearly complete prompting across the material demonstrates how iterative collaboration and backward design process were successfully used to intentionally strengthen the scaffolding necessary for students to engage fully with complex scientific practices. For example, in the case of SP6, constructing explanations and engaging in arguments from evidence, the partially aligned μCUREs all had the first three criteria:

> Criteria 1: Laboratory experiment gives an event, observation, or phenomenon.
> Criteria 2: Laboratory experiment asks students to make a claim based on the given event, observation, or phenomenon.
> Criteria 3: Laboratory experiment asks students to provide evidence in the form of data or observations to support the claim.

However, during Iterative Collaboration 3, the teaching materials were revised in such a way that they now included the final Criteria 4 (Laboratory experiment asks students to provide reasoning about why the evidence supports the claim), moving SP6 to the "complete prompting" category.

Conclusions

The combined analysis of instructional materials and student perceptions strongly supports the conclusion that

TABLE 4. Level of Inquiry for Partially and Fully Aligned μCUREs

Partially aligned μCUREs				
	Institution 1	Institution 2	Institution 3	Institution 4
Laboratory characteristic	Subcode			
Problem/question	Partially provided	Partially provided	Partially provided	Provided
Theory/background	Not provided	Not provided	Not provided	Provided
Procedure/design	Not provided	Not provided	Not provided	Partially provided
Results analysis	Not provided	Not provided	Not provided	Not provided
Results communication	Not provided	Not provided	Not provided	Not provided
Conclusions	Not provided	Not provided	Not provided	Not provided
Overall level of inquiry	Level 2.5: nearing authentic inquiry	Level 2.5: nearing authentic inquiry	Level 2.5: nearing authentic inquiry	Level 1.5: nearing open inquiry

Fully aligned published μCUREs	
Laboratory characteristic	Subcode
Problem/question	Partially provided
Theory/background	Not provided
Procedure/design	Not provided
Results analysis	Not provided
Results communication	Not provided
Conclusions	Not provided
Overall level of inquiry	Level 2.5: Nearing authentic inquiry

the backward design process using practical tools (the backward design modules and assignments) successfully led to the development of μCUREs that integrated authentic research with necessary instructional scaffolding. The modified inquiry rubric results confirmed that the materials consistently fostered experiences categorized as "nearing authentic inquiry," indicating that the design placed substantial decision-making autonomy on the students, a core component of authentic research. This high measure of authentic inquiry aligns strongly with student perceptions captured by the STIR data, which consistently demonstrated a stable, learner-centered environment across both the pilot and partially aligned implementation years, confirmed by the lack of significant differences across any rubric category. However, a more nuanced understanding of the μCUREs emerged when comparing the levels of prompting for scientific practices across implementations. Specifically, the partially aligned μCUREs often missed the final critical criteria, requiring justification in scientific practices. The subsequent iterative collaboration, driven by the backward design methodology, rectified this critical scaffolding gap in the fully aligned μCURE, leading to "complete prompting" for five out of six intended scientific practices. Therefore the combined results affirm that the backward design process successfully produced high quality CUREs by identifying and strengthening essential scaffolding through iterative collaboration anchored in curriculum design theory.

Implications and Future Work

A primary implication arising from this study is the strong suggestion that the backward design framework is highly effective, especially when implemented within a professional development format. Successful implementation of this framework appears to rely on two crucial strategies: the active utilization of curriculum design theory through practical tools, and sustained collaborative iteration. It is suggested that instructors who wish to develop high-quality CUREs employ backward design and utilize practical guidance, such as the tools provided in the supplemental

TABLE 5. Scientific Practices in Partially and Fully Aligned μCUREs

	SP3	SP4	SP5	SP6	SP7	SP8
	Planning/ carrying out investigations	Analyzing/ interpreting data	Using math and computational thinking	Construct explanation and engage in argument from evidence	Evaluate information	Defining problems and designing solutions
Partially aligned μCUREs						
Institution 1	C	C	T	T	C	P
Institution 2	C	C	C	T	C	T
Institution 3	C	C	T	T	C	T
Institution 4	C	C	T	T	C	T
Fully aligned μCUREs						
	C	C	C	C	C	T

Note: C, complete prompting; P, partial task; T, task only.

information (linked in Figure 1). Furthermore, integrating a strong foundation of collaborative iteration is encouraged, perhaps by finding a chemical education collaborator or recruiting other disciplinary experts who can work together to build the CURE.

Looking ahead, future scholarly work could benefit from integrating the scientific practice criteria identified by Van Wyk et al. (2025) earlier in the backward design process. Additionally, obtaining specific feedback regarding the presence of scientific practices at various collaborative iteration checkpoints might offer instructors valuable data points that can inform their pedagogical decisions. Although the evaluation procedures (such as the Van Wyk 2024 aand 2025 publications) were published during the grant's execution, limiting their initial inclusion, integrating SP presence feedback from the outset could maximize the effectiveness of the backward design process.

Limitations

A major limitation acknowledged in this work is the inherently small sample size, which consisted solely of the analytical co-principal investigators involved in the development and implementation of the μCURE and their students. Additionally, although the iterative nature of the backward design process is a central theme, a complete understanding of this progression was constrained because the final implementation of the fully aligned μCURE was not included in the student perception data; the findings were brought to the community without waiting the additional year that would have been required to collect and analyze data from that final implementation.

Data Availability Statement

The data reported in this manuscript are available upon reasonable request.

Institutional Review Board

All research efforts described here were approved under appropriate IRB protocol granted by each institution involved.

Conflict of Interest

No conflicts of interest to declare.

Acknowledgments

This material is based upon work supported by a National Science Foundation DUE-IUSE grant (award 2215768). Thank you for the research support provided by undergraduate research assistants at St. Mary's College of Maryland: Afua Atta-Poku, Nia Delauter, and Emile Walker. Also, great appreciation goes to the students who completed the μCURE labs in these initial implementations.

References

Bodzin, Alec M., and Karen M. Beerer. 2003. "Promoting Inquiry-Based Science Instruction: The Validation of the Science Teacher Inquiry Rubric (STIR)." *Journal of Elementary Science Education* 15(2): 39–49. doi: 10.1007/BF03173842

Bruck, Laura B., Stacey Lowery Bretz, and Marcy H. Towns. 2008. "Characterizing the Level of Inquiry in the Undergraduate Laboratory." *Journal of College Science Teaching* 38(1): 52–58.

Carter, Frances D., Marvin Mandell, and Kenneth I. Maton. 2009. "The Influence of On-Campus, Academic Year Undergraduate

Research on STEM PhD Outcomes: Evidence from the Meyerhoff Scholarship Program." *Educational Evaluation and Policy Analysis* 31: 441–462. doi: 10.3102/0162373709348584

Chang, Joyce C. 2002. *Women and Minorities in the Science, Mathematics and Engineering Pipeline*. Los Angeles: ERIC Clearinghouse for Community Colleges

Cooper, Katelyn M., Paula A. G. Soneral, and Sara E. Brownell. 2017. "Define Your Goals Before You Design a CURE: A Call to Use Backward Design in Planning Course-Based Undergraduate Research Experiences." *Journal of Microbiology and Biology Education* 18(2): 1–7. doi: 10.1128/jmbe.v18i2.1287

Corwin, Lisa A., Sandra L. Laursen, Janet L. Branchaw, Kevin Eagan, Mark J. Graham, David I. Hanauer, Gwendolyn A. Lawrie, et al. 2014. "Assessment of Course-Based Undergraduate Research Experiences: A Meeting Report." *CBE–Life Sciences Education* 13: 29–40. doi: 0.1187/cbe.14-01-0004

Corwin, Lisa A., Christopher R. Runyon, Eman Ghanem, Moriah Sandy, Greg Clark, Gregory C. Palmer, Stuart Reichler, Stacia E. Rodenbusch, and Erin L. Dolan. 2018. "Effects of Discovery, Iteration, and Collaboration in Laboratory Courses on Undergraduates' Research Career Intentions Fully Mediated by Student Ownership." *CBE–Life Sciences Education* 17(2): ar20. doi: 10.1187/cbe.17-07-0141

Dabney-Smith, Valschkia Lisette. 2009. "A Multi-Level Case Study Analysis of Campus-Based Male Initiatives, Programs, and Practices and the Impact of Participation on the Perceptions of First-Year African American Male Community College Students in Texas." PhD diss., University of Texas, Austin.

Dolan, Erin, and Gabriela Weaver. 2021. *A Guide to Course-Based Undergraduate Research*. Macmillan Higher Education.

Eagan, M. Kevin, Sylvia Hurtado, Mitchell J. Chang, Gina A. Garcia, Felisha A. Herrera, and Juan C. Garibay. 2013. "Making a Difference in Science Education: The Impact of Undergraduate Research Programs." *American Educational Research Journal* 50: 683–713. doi: 10.3102/0002831213482038

Enhancing Learning by Improving Process Skills in STEM. 2025. ELIPSS.com. https://elipss.com/index.html

Frederick, Kimberley A., Maury E. Howard, Kelly Y. Neiles, Daniel F. Scott, and Rebecca A. Hunter. 2025. "Development of a Safe, Scalable, Course-Based Undergraduate Research Experience for Analytical Chemistry: The μCURE Project." *Journal of Chemical Education* 102: 4024–4032. doi: 10.1021/acs.jchemed.5c00809

Gwet, Kilem. 2002. "Kappa Statistic Is Not Satisfactory for Assessing the Extent of Agreement between Raters." *Statistical Methods for Inter-Rater Reliability Assessment* 1(6): 1–6.

Harrison, Melissa, Deb Dunbar, Lana Ratmansky, Kelly Boyd, and David Lopatto. 2011. "Classroom-Based Science Research at the Introductory Level: Changes in Career Choices and Attitude." *CBE–Life Sciences Education* 10: 279–286. doi: 10.1187/cbe.10-12-0151

Hunter, Anne, Sandra Laursen, and Elaine Seymour. 2007. "Becoming a Scientist: The Role of Undergraduate Research in Students' Cognitive, Personal, and Professional Development." *Science Education* 91: 36–74. doi: 10.1002/sce.20173

Laverty, James T., Sonia M. Underwood, Rebecca L. Matz, Lynmarie A. Posey, Justin H. Carmel, Marcus D. Caballero, Cori L. Fata-Hartley, Diana Ebert-May, Sarah E. Jardeleza, and Melanie M. Cooper. 2016. "Characterizing College Science Assessments: The Three-Dimensional Learning Assessment Protocol." *PLoS One* 11(9): e0162333. doi: 10.1371/journal.pone.0162333

Lopatto, David. 2007. "Undergraduate Research Experiences Support Science Career Decisions and Active Learning." *CBE–Life Sciences Education* 6: 297–306. doi: 10.1187/cbe.07-06-0039

McHugh, Mary L. 2012. "Interrater Reliability: The Kappa Statistic." *Biochemia Medica* 22: 276–282. doi: 10.11613/BM.2012.031

National Research Council. 2012. *A Framework for K–12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: National Academies Press.

Neiles, Kelly Y., and Katy Arnett. 2021. "Backward Design of Chemistry Laboratories: A Primer." *Journal of Chemical Education* 98: 2829–2839. doi: 10.1021/acs.jchemed.1c00443

Provost, Joseph J. 2022. "Developing Course Undergraduate Research Experiences (CUREs) in Chemistry." *Journal of Chemical Education* 99: 3842–3848. doi: 10.1021/acs.jchemed.2c00390

Provost, Joseph J., Jessica K. Bell, and John E. Bell. 2019. "Development and Use of CUREs in Biochemistry." American Chemical Society *ACS Symposium Series* 1337: 143–171. doi: 10.1021/bk-2019-1337.ch007

Reynders, Gil, Erica Suh, Renée S. Cole, and Rebecca L. Sansom. 2019. "Developing Student Process Skills in a General Chemistry Laboratory." *Journal of Chemical Education* 96: 2109–2119. doi: 10.1021/acs.jchemed.9b00441

Rowland, Susan L., Gwen A. Lawrie, James B. Y. H. Behrendorff, and Elizabeth M. J. Gillam. 2012. "Is the Undergraduate Research Experience (URE) Always Best? The Power of Choice in a Bifurcated Practical Stream for a Large Introductory Biochemistry Class." *Biochemistry and Molecular Biology Education* 40: 46–52. doi: 10.1002/bmb.20576

Russell, Susan H., Mary P. Hancock, and Jennifer McCullough. 2007. "Benefits of Undergraduate Research Experiences." *Science* 316: 548–549. doi: 10.1126/science.1140384

Sadler, Troy D., and Laura McKinney. 2010. "Scientific Research for Undergraduate Students: A Review of the Literature." *Journal of College Science Teaching* 39(5): 43–49.

Van Wyk, Andrea L., Ardith Bhinu, Kimberley A. Frederick, Marya Lieberman, and Renée S. Cole. 2025. "Bridging the Science Practices Gap: Analyzing Laboratory Materials for Their Opportunities for Engagement in Science Practices." *Journal of Chemical Education* 102: 970–983. doi: 10.1021/acs.jchemed.4c00744

Van Wyk, Andrea L., Kimberley A. Frederick, Marya Lieberman, and Renee S. Cole. 2024. "Increasing Authenticity of the Laboratory through the MICRO Project: Analysis of Analytical Chemistry Laboratory Experiments for Their Level of Inquiry." *Journal of Chemical Education* 102: 3–14. doi: 10.1021/acs.jchemed.3c00945

Weaver, Gabriela C., Cianán B. Russell, and Donald J. Wink. 2008. "Inquiry-Based and Research-Based Laboratory Pedagogies in Undergraduate Science." *Nature Chemical Biology* 4: 577–580. doi: 10.1038/NCHEMBIO1008-577

Wiggins, Grant, and Jay McTighe. 2011. *The Understanding by Design Guide to Creating High-Quality Units*. Alexandria, VA: ASCD.

Kelly Y. Neiles
Mary's College of Maryland, kyneiles@smcm.edu

Kelly Y. Neiles is a professor of chemistry and associate dean of faculty at St. Mary's College of Maryland. Her research is focused on large scale curricular reform of STEM courses with a focus on introductory and laboratory classes.

Rebecca A. Hunter is an associate professor in the chemistry department at the College of New Jersey. Her research focuses on analytical method development, including novel electrochemical sensors, paper-based analytical devices, and chromatographic methods.

Maury E. Howard is a professor in the chemistry and biochemistry department at Virginia Wesleyan University. Her research focuses on analytical method development, including atomic spectroscopy, chromatography, and paper-based analytical devices, primarily for environmental and pharmaceutical analysis applications.

Daniel F. Scott is an associate professor of chemistry at Centre College. His research focuses on the development and optimization of low-cost, point-of-care diagnostics with applications for human health, environmental science, and food and beverage manufacturing.

Kimberley A. Frederick is the Charles Lubin Family Chair for Women in Science at Skidmore College. Her research focuses on development of colorimetric microfluidic paper analytical devices for applications in environmental, pharmaceutical, and biomedical analyses.