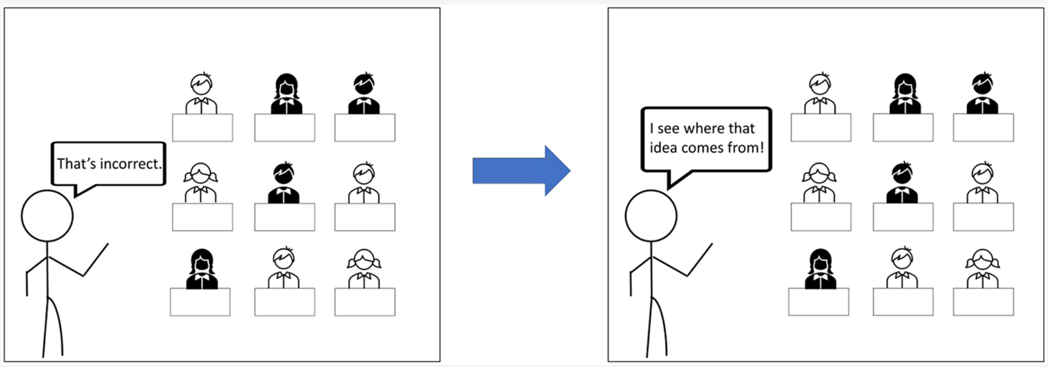
JOURNAL OF CHEMICAL EDUCATION

pubs.acs.org/jchemeduc Article

Investigating How Teachers' Formative Assessment Practices Change Across a Year

Timothy N. Abell* and Hannah Sevian

Cite This: J. Chem. Educ. 2021, 98, 2799–2808 Read Online



ABSTRACT: Teaching chemistry as a practice rather than as a mere collection of facts demands that teachers modify their practices, particularly their approach to formative assessment (FA). In this study, we investigated how teachers' FA practices changed as a result of their participation in a professional development program designed with a Chemical Thinking perspective. Four FA portfolio chapters were collected from 19 secondary school teachers over the course of a year. The analysis of the FA portfolio chapters gave insight into changes in teachers' FA practice in the areas of task design, purpose, and focus when evaluating student work. All teachers implemented changes in at least one of these areas of analysis, with about half of them creating FA tasks that demonstrated changes in all three dimensions. Changes in FA design were the most prevalent among participating teachers, shifting from tasks designed to explore the acquisition of knowledge to tasks that explored student reasoning with chemistry. On the other hand, changes in evaluation focus were the least common as most teachers centered their attention on the correctness of students' answers rather than on the nature of students' chemistry ideas. The results of our investigation point to areas in which chemistry teachers require substantial support to effectively use FA in their classrooms.

KEYWORDS: *Elementary/Middle School Science, High School/Introductory Chemistry, Professional Development, Chemical Education Research, Testing/Assessment*

FEATURE: Chemical Education Research

■ INTRODUCTION

The Next Generation Science Standards emphasize the need for teachers to focus instruction on core ideas, cross-cutting concepts, and science practices rather than on the memorization of vast amounts of content knowledge.[1,2] It is expected that teachers will create many opportunities for students to make sense of phenomena in their surroundings and integrate their understanding with their everyday lives. To support this type of learning, teachers need to modify traditional practices to support students' meaningful engagement with scientific ideas and ways of thinking and acting. It is particularly important that teachers pay attention to the substance of student thinking, critically reflect on students' ideas, and act responsively to advance student understanding.[3] This way of teaching thus demands a high level of expertise in formative assessment (FA) and a deep understanding of chemistry.

According to Bell and Cowie, FA is "the process used by teachers and students to recognize and respond to student learning in order to enhance that learning, during the learning."[4] Tools that assist in the FA process include quizzes, worksheets, or hands-on laboratory activities, and other activities that provide opportunities for teachers to gain insight into students' thinking or understanding. FA has been found to have a major impact on students' learning, and therefore it has been identified as a high-leverage teaching practice.[5,6] The

Received: April 6, 2021
Revised: July 8, 2021
Published: August 17, 2021

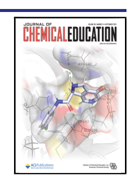


Table 1. Participating Teachers' Demographic Information

Teacher	Years Teaching	Primary Subject	Grade Level	Other Subjects Taught	Type of School
A	6	Chemistry	10, 11, 12	Physics	Large comprehensive urban high school
B	5	Chemistry	10	None	Urban high school for English learners
D	11	Chemistry	11, 12	Earth and Environmental Science	Large selective urban high school
E	3	Chemistry	11, 12	Biology, Earth, and Environmental Science	Small learning community urban high school
F	6	Chemistry, AP Chemistry	10, 11, 12	Biology	Large comprehensive urban high school
G	12	General Science	6, 7, 8	None	Spanish bilingual urban middle school
I	24	Chemical Interactions	8	General Science	Large inclusive urban middle school
J	18	Chemistry, AP Chemistry	10, 11, 12	Physics	Large comprehensive suburban high school
K	11	Chemistry, AP Chemistry	10, 11, 12	Physics	Large comprehensive suburban high school
L	3	Chemistry	9, 10, 11, 12	General Science, Forensic Science	Large comprehensive urban high school
M	7	Chemistry	10, 11	Biology, Biotechnology	Large selective urban high school
N	5	Chemistry, AP Chemistry, Inclusion Chemistry	10, 11, 12	None	Large comprehensive urban high school
O	8	Chemistry, AP Chemistry	10, 11	Biology, General Science, Forensic Science	Large comprehensive suburban high school
P	3	Biology, AP Biology	11, 12	Chemistry	Vocational-technical urban high school
Q	6	Chemistry	9, 10, 11, 12	None	Arts magnet urban high school
R	7	Chemistry, AP Chemistry	10, 11, 12	Earth and Environmental Science	Large selective urban high school
S	3	Chemistry	11, 12	None	Small learning community urban high school
T	14	Chemistry, AP Chemistry	11, 12	Physics	Small learning community urban high school
U	5	General Science	6, 7, 8	Biology, Physics, Earth and Environmental Science, Technology/Engineering	Spanish bilingual urban middle school

formal FA process involves setting a goal or purpose for the assessment, designing a task to achieve the established goal, implementing the task in the classroom, and analyzing the information generated to determine how to best proceed.[4,7]

The design of effective FA tasks is a critical step in the process as it determines the insights that can be gained into students' knowledge and thinking.[7−10] FA tasks that are not aligned with learning goals can limit teachers' ability to draw meaningful conclusions about student understanding.[8] Use of FA tasks that ask students to make connections between concepts or build explanations tends to be indicative of teaching expertise in this area.[9] Of the four phases of FA—design, launch, implementation, and evaluation—the design phase involves the highest intellectual demand, and limits the level of intellectual demand that follows.[10]

Kloser et al. identify two aspects of teachers' FA practices that most reliably indicate teacher effectiveness: (1) the extent to which FA tasks explore students' ability to understand complex scientific concepts and build connections among them, and (2) the extent to which FA tasks require students to explain or justify their reasoning using scientific concepts and evidence.[9] When setting goals for FA, less effective teachers tend to focus their efforts on evaluating whether students can produce or identify correct answers to closed-response questions. Their analysis of students' answers is often limited to characterizing the percent of students who get the right or wrong answers.[11−15] Even when analyzing students' written responses to open questions, many teachers adopt an evaluative approach, focusing on the correctness of expressed ideas rather than trying to make sense of student reasoning to come up with interventions that advance student understanding.[15−19] More effective teachers often recognize productive cognitive and epistemic resources when evaluating student work that can be leveraged to support learning.[14−16,20]

Previous research has focused on characterizing teachers' FA practices at set points in their training or careers. In this investigation we seek to contribute to the field by characterizing how chemistry teachers' FA practices change over time while participating in a professional development program focused on developing such practices. In our analysis, we identified and characterized the extent to which different aspects of participating teachers' FA practices changed over a year. Our work provides insights into areas in which chemistry teachers are likely to require substantial support to effectively use FA to uncover students' underlying thinking about chemistry and advance student learning.

■ FOCUS ON CHEMICAL THINKING

With the highly contextual nature of chemistry teaching and the diversity of participating teachers' situations in mind, the professional development (PD) program was designed to

https://doi.org/10.1021/acs.jchemed.1c00356
J. Chem. Educ. 2021, 98, 2799−2808

support teachers in adopting a "Chemical Thinking" (CT) perspective in their courses.[21] Assessment goals and practices are tightly linked to the nature of the curriculum that is followed. Participating teachers in our study worked in different schools and taught a variety of chemistry and science classes. While the PD involved domain-general aspects of FA, teachers were engaged in designing and enacting FA practices related to chemistry content specifically. Within the CT perspective, chemistry is conceptualized as a productive way of thinking that can be used to make sense of and address relevant problems in our surroundings rather than as a static body of knowledge to be learned. Developing chemical thinking is expected to help students make informed decisions, build justifications, design solutions, test conjectures, and evaluate outcomes using chemical ways of knowing, thinking, and acting. The CT framework highlights six core crosscutting ways of reasoning to be developed in the chemistry classroom: chemical identity,[22,23] structure−property relationships,[24,25] chemical causality,[26,27] chemical mechanism,[26,27] chemical control,[14] and benefits−costs-risks.[28,29]

Teachers who adopt a CT perspective in their teaching create opportunities for students to apply their knowledge and experiences in authentic tasks that demand the analysis, synthesis, or transformation of matter for relevant purposes. As students work on these tasks, teachers engage in FA practices that seek to identify productive cognitive and epistemic resources that students express or manifest and how those resources can be leveraged to achieve task goals and advance student understanding. Teachers' attention is directed at students' reasoning while adopting an interpretive, rather than evaluative, stance. Teachers' decisions are geared toward strengthening the six crosscutting ways of chemical thinking rather than the ability to reproduce normative answers. The CT framework was used in this study to analyze changes in the FA practices of teachers participating in the PD program. It is within this framework that we sought to characterize changes in participating teachers' FA practices.

■ RESEARCH QUESTION

This investigation was guided by the following research question:

- What aspects of teachers' chemistry FA practices in the areas of task design, purpose, and evaluation focus change as they participate in a one-year professional development program guided by the CT framework?

■ RESEARCH METHODS

Context and Participants

This study focused on teachers (N = 19) participating in a yearlong PD program. Teachers were selected for PD through a double-blind review of applications conducted by teachers and researchers on the project team. The teachers' years of experience, school, and teaching situation were considered to diversify selected participants. To qualify for the program, teachers needed to teach middle school science or high school chemistry in a specific urban New England area, have at least three years of teaching experience, and obtain approval from their principal. Data were collected from three cohorts in 2017−2018 (n = 5), 2018−2019 (n = 7), and 2019−2020 (n = 7). The 19 teachers in this study were either middle school science (n = 3), high school chemistry (n = 15), or high school biology (n = 1) teachers (when applying to the program, the high school biology teacher was scheduled to teach chemistry before being reassigned after being accepted). As summarized in Table 1, the teachers had varying years of teaching experience, taught in different school settings, and taught various other subjects. However, the sample size is not large enough to draw conclusions based on any demographic factors. The university's IRB and the school districts' research compliance offices approved the study. All teachers provided consent to participate and were compensated with either graduate level credits in chemistry or a monetary stipend. Consent by students in these teachers' classes was obtained directly (students over 18 years old) or from a parent or guardian.

The PD was designed to develop teachers' FA practices in three areas: the design and selection of FA tasks, the evaluation of student work, and the actions to take in response to the results of the assessment. The PD sought to help teachers shift from using FA to differentiate students based on performance levels to using FA to elicit and advance students' chemical thinking. When evaluating student work, the PD aimed to shift teachers' focus from an evaluative stance (marking students as right and wrong) to an interpretive stance (interpreting student thinking and identifying productive resources). Lastly, the PD emphasized teachers' responsive actions directed at advancing student thinking rather than simply correcting their errors and misconceptions. The PD occurred once a month, for ten months, with each meeting lasting 3 h. The teachers participated in activities that helped them to understand the design features of an FA that make it both accessible for all students and able to reveal students' chemical thinking.[30] The teachers also engaged in a CT-oriented FA that was designed by the PD leadership team. As the teachers designed their own FA, they were able to brainstorm with and get feedback from the PD facilitators. The teachers also took part in three focus groups, during which teachers shared an FA task along with student work generated by that task. The teacher, along with their peers in the focus group, would critique the task and discuss how they would evaluate the students' work and what feedback they could provide. This activity allowed the teachers to apply what they were learning about CT to their own FA designs and evaluation styles.

Data Collection

Data were collected in the form of four FA portfolio chapters generated by each participating teacher over the course of a year. The first portfolio chapter was submitted as a part of the teachers' application to the PD (portfolio A), the second submitted three months into the PD (portfolio 1), the third submitted seven months into the PD (portfolio 2), and the final portfolio was submitted toward the end of the school year (corresponding to the end of the PD) (portfolio 3). The FA portfolio chapters were modeled after Scoop Notebooks,[31−33] which have been used to evaluate science and mathematics teaching practices. For each portfolio chapter the teachers were asked to include a copy of an FA task they had implemented in their class, their purpose for using it, a description of where the FA fit in their curriculum, an expected answer from a proficient student, and descriptions of confusions or misunderstandings the teacher expected to see in students' responses. The teachers also provided deidentified work from three consenting students that illustrated the range of answers received and the teacher's evaluation of this work along with hypothetical

https://doi.org/10.1021/acs.jchemed.1c00356
J. Chem. Educ. 2021, 98, 2799−2808

Table 2. Coding Scheme Definitions

Category	Definition	Examples
	Design	
Tool-oriented	Focuses on developing a specific skill or targeting a specific content knowledge using questions that have one correct answer.	Students are given three solutions of NaCl at different concentrations and asked to determine which is the most concentrated and which contains the most moles of solute.
Correctness-assessing structure	Tasks that have the potential to focus on students' thinking related to a chemical thinking strand, but the structure of the questions does not give students opportunities to share their thinking about the task. These questions often have one correct answer,	Students are asked to use cards (H_2O, H^+, OH^-) to represent what happens when an acid is added to pure water. The students are then asked closed questions with one correct answer (e.g., "What happens to the amount of H^+ in the solution?").
Thinking constraining	Tasks that present a limited number of outcomes or focus students' attention on a limited number of factors. Within these constrains students are given the opportunity to share their thinking about one or more chemical thinking strands.	Students are asked to comment on the correctness of the particulate representations of water after it has evaporated.
Chemical Thinking-oriented	Tasks that present questions that chemistry is uniquely positioned to answer. This usually rests in the explanation of a phenomenon, often familiar or occurring in the real world, and asks students to apply their content knowledge to explain the phenomenon.	Students are tasked with using their knowledge to determine why a smaller person died while a larger person lived after consuming the same amount of caffeine.
	Purpose	
Conceptual mastery	Assessing students' understanding of concepts or mastery of skills.	"The central goal of this assessment is the concept of concentration. The concept of concentration is very important in chemistry and all its applications."
Conceptual mastery and chemical thinking	Assessing students' understanding of concepts or mastery of skills and uncovering students' reasoning about a chemistry-related phenomenon.	"Highlight a few key concepts and [see] what students are thinking about solutions for this problem."
Chemical thinking	Uncovering students' reasoning about a chemistry-related phenomenon.	"I was curious to know their thoughts on what makes the three states of matter different from one another."
	Evaluation Focus	
Conceptual understanding	Focused on what students did right or wrong, what was missing from students' work, and what the students could do to provide the correct answers.	"[I see] understanding of runoff and infiltration of impervious and pervious land cover, respectively."
Conceptual understanding and students' thinking	Focused on what students did right or wrong, what was missing from students' work, and what the students could do to provide the correct answers. These evaluations also included feedback to either further elicit or advance students' thinking.	"She was able to change the amount of molecules in her drawing based on the number of spoonfuls. The only thing I didn't see is language centered around the specificity of concentration. Why did the molecules stay the same? What happened specifically to the taste? Why did the taste change?"
Students' chemical thinking	Focused on making inferences about the thinking and reasoning underlying students' responses.	"I notice that this student, as many did, had the idea that acids "destroy" and "melt" things. I did notice, with her and with many other students, that on the exit slip they wrote things like "break down" that sound more like they are talking about a chemical reaction than "destroying" matter."

https://doi.org/10.1021/acs.jchemed.1c00356
J. Chem. Educ. 2021, 98, 2799−2808

feedback that they would supply to the students if given the opportunity. For the final portfolio chapter, the teachers were asked to design an FA aligned with the CT perspective. Teachers received guidance from the facilitators and their peers as they worked on this assignment. The expectations for the portfolio chapters remained the same throughout the yearlong PD. The template for a FA portfolio chapter, together with an example supplied to the teachers, can be found in the Supporting Information.

Data Analysis

Each FA portfolio chapter was analyzed in relation to three key elements: (1) FA task design, (2) FA stated purpose, and (3) teachers' focus when evaluating student work. These analyses were completed by adapting a coding scheme developed in a previous study and summarized in Table 2.[20]

From the perspective of task design, the FA activities were analyzed in terms of their potential to elicit students' ability to productively engage in chemical thinking, from low (tool-oriented) to high (CT-oriented). Tool-oriented FA tasks target specific skills or knowledge and have prescribed correct answers; for example, assessing students' understanding of concentration by determining their ability to apply the concentration formula $C = n/V$. Correctness-assessing FA tasks reveal more about students' ways of explaining, but there is an expected normative response. A correctness-assessing FA task might, for example, ask students to simulate the mechanism of what happens when an acid is added to water using cards (H_2O, OH^-, H^+), but only ask them to state the trends they notice (How does the amount of hydrogen ions change?). These types of questions only allow the students to give a one-word answer (increase or decrease) and does not give the teacher insight into the students' thinking.

FA tasks that are thinking constraining open opportunities for students to express their ideas but within a set of specified options. For example, a teacher might ask students about the evaporation of water on the particulate level, but rather than asking the students to generate their own representations, the teacher asks the students to comment on the correctness of provided representations. Finally, CT-oriented FA tasks allow students to apply chemical thinking in more open ways to build explanations, make decisions, or solve problems. An example of a CT-oriented task is one in which students are presented with a scenario of someone dying from caffeine ingestion while another, larger person did not die from ingesting the same amount of caffeine. The students are then tasked with applying their knowledge of chemistry and toxins to determine why this may have occurred (chemical mechanism and causality).

Teachers' stated purposes for the FA tasks were characterized based on the answers provided to several prompts included in their portfolio chapters, which asked teachers to describe "what the central idea or concept being targeted was, where the FA was positioned in the curriculum, the benefit or "special power" of the assessment, and expected answers from proficient students and confusing aspects students are likely to encounter"[20] Narratives were developed for each portfolio chapter based on the answers provided and used to classify the purpose of an FA task into one of three categories (Table 2): (1) conceptual mastery, (2) chemical thinking, or (3) both. When including a task with a conceptual mastery goal, teachers often stated the concept that they were assessing and why it was important for students to understand it. When the purpose was to uncover chemical thinking, teachers expressed interest in eliciting students' "thoughts" or "ideas" about why and how chemical phenomena happened. In some cases, teachers included both types of goals in their rationales.

The focus of teachers' evaluation was characterized based on what teachers noticed and interpreted from the pieces of student work included in a portfolio chapter and the type of hypothetical feedback they provided. Participating teachers focused on (1) normative conceptual understanding, (2) chemical thinking, or (3) both. When focused on normative understanding, teachers tended to describe what students wrote, identify what concepts the students understood, and provide feedback that guided students toward the correct answer. When focused on chemical thinking, teachers highlighted productive or problematic student ideas, seeking to make sense of student reasoning regardless of the correctness of an answer. Teachers who focused on both normative understanding and chemical thinking tended to attend to students' correct or incorrect ideas, but their feedback included questions directed at further eliciting students' thinking or to advancing their reasoning.

Changes in FA design, purpose, and evaluation focus were identified by comparing the results of the analysis across portfolio chapters created by teachers at different moments of the PD. Qualitative graphs were built to facilitate the identification of patterns of change in the sample (see Supporting Information). Several areas of trustworthiness were considered during the analysis of all data: dependability, credibility, and transferability.[34] The dependability of the data was ensured through thorough documentation of the data collection and analysis processes. The credibility of the results was established through a second rater, a graduate student not affiliated with the data collection or initial coding, who coded all four portfolio chapters from two teachers. Coding between the two raters was the same except for a few cases in which discussions were held until agreement was reached. The diversity of the participants (Table 1) supports the transferability of the results and conclusions drawn from this study.

■ RESULTS

In this section we summarize how participation in the PD program affected teachers' FA practices in the areas of task design, purpose, and evaluation focus. None of the participating teachers consistently progressed toward a CT orientation in all three areas during the PD. Some of them (n = 9) exhibited changes in all categories of analysis, but these developments were not necessarily consistent or aligned across design, purpose, and evaluation focus, and these changes did not always occur toward a more CT orientation. Other participants (n = 8) exhibited variations in FA practice in only two areas, while a minority (n = 2) manifested variation in only one dimension. All teachers (n = 19) exhibited variations in FA task design, with 12 of them submitting CT-oriented tasks at some point during the PD year and nine teachers including that type of task in their final portfolio chapter. Most participating teachers (n = 17) also varied in their stated purpose for using FA tasks in the classroom. At the beginning of the PD, many of these teachers used FA with the main goal of assessing content mastery, but 14 of them ended the PD year using FA tasks to uncover students' ability to apply chemical thinking. Only nine teachers demonstrated changes in their evaluation focus and three of them manifested increased focus on students' chemical thinking in the analysis

https://doi.org/10.1021/acs.jchemed.1c00356
J. Chem. Educ. 2021, 98, 2799−2808

of FA products by the end of the PD. The remaining six teachers who manifested change in this area maintained a focus on their students' normative conceptual understanding but began to provide feedback directed at eliciting or advancing student thinking. These general results are summarized in Table 3.

Table 3. FA Practice Changes by Area

Category	Teachers Who Demonstrated Change	Teachers Who Manifested a CT Perspective in Their Last Portfolio Chapter
Design	19	9
Purpose	17	6
Evaluation	9	3

While all participants in the PD made changes in at least one area of their FA practice, there was a wide range of variation with regard to how and in what areas each of them progressed. In the following sections we seek to better characterize this variation.

Change in One Area: Task Design

Only two of the study participants manifested variation in a single target area during the PD year, and this category was FA task design. This change, however, was not consistent. One of these teachers, Teacher I, provided two FA tasks that were tool-oriented (portfolios A and 2) and one characterized as correctness-assessing (portfolio 1). The last task (portfolio 3) was more open but at the level of thinking constraining. This teacher recognized the value of designing FA tasks that more openly elicited students' ideas but felt that these activities were more difficult for students to complete and for teachers to manage:

"You know, I want to give kids the opportunity to show me what they're thinking in a, other than just choosing a multiple-choice answer. Cause there's guessing involved, and a lot of times they'll just write anything down. If they have to explain their thinking, then maybe it's a little bit more authentic. But then, you know, students find open-ended things very challenging and very threatening, and especially kids that have limited literacy skills. And so then it's often a battle. You didn't answer that one. Let's go back and try that, and they're very resistant to doing more open-ended things as a general rule." (Teacher I)

Despite these concerns, this teacher's last FA task moved beyond using fact-oriented multiple-choice questions by asking students to describe in their own words the characteristics of different states of matter and the main differences between them. The teacher, however, still made available a "science vocabulary word bank" to support students in writing their descriptions, designed/approached the task with the main purpose of evaluating students' comprehension of targeted concepts, and focused on the correctness of responses when evaluating student work. This teacher was more focused on literacy and how the structure of the assessment affects how students answer it, an orientation which is characteristic of teachers who take a tool-oriented approach to chemistry. This tool-oriented approach was evidence that teachers were not connecting the way they were assessing (i.e., literacy) to the nature of what was being assessed (i.e., chemistry).

The second teacher in this category, Teacher P, was new to using formal structured FA tasks to assess student understanding and was used to posing "*one quick question*" to determine "*what [students] know or don't know.*" This teacher relied on "*on-the-spot eliciting*" during a class to gauge knowledge comprehension. This teacher started the PD program providing a task that was thinking constraining, but then included two CT-oriented activities in portfolios 1 and 2, although they maintained a content mastery focus and an evaluative stance while analyzing student work. The CT orientation, however, was not sustained in the last portfolio, where the associated task was tool-oriented.

Change in Two Areas: Task Design and Purpose

Eight of the participating teachers exhibited variations in the areas of FA task design and purpose, but not necessarily in an aligned or consistent manner (Figure 1). Six of them submitted

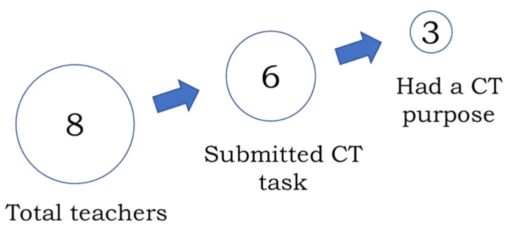


Figure 1. Total number of teachers who made changes to their design and purpose and the subset of teachers who had a CT focus in those areas in at least one FA portfolio.

one CT-oriented FA at some point during the PD year, but only three out of eight teachers in this group indicated at least once that the main goal of their FA task was to uncover chemical thinking. Only one of these three teachers, Teacher B, included a FA task that was CT-oriented both in design and in purpose. Nevertheless, toward the end of the PD year, five out of the eight teachers provided FA tasks that were progressing toward a CT perspective in these two areas. Their FA task designs included more open questions that more frequently asked students to explain real-world phenomena. There was also a shift in the stated purpose of the assessment. For example, when describing goals of the FA task included in portfolio 2, Teacher B stated that the purpose of the FA was to "*teach students about atomic structure and valence electrons,*" which revealed a central focus on evaluating conceptual understanding. Contrastingly, with the FA task included in portfolio 3, this teacher sought to "*learn about what students already know. . .reveal what [students] believe to be the main differences [between ice, water, and steam] . . .and indicate whether students are thinking about macro-level properties or micro-level properties.*" This statement illustrates a shift in purpose toward eliciting students' prior knowledge, ideas, and ways of reasoning with chemistry (chemical identity).

While all these teachers made changes in their design and purpose, how and when the design and purpose of their submitted tasks changed were quite varied (Figure 2). A subset of these teachers (Teacher B, Teacher F, Teacher O, Teacher U) changed their FA design and purpose simultaneously, within the same FA portfolio chapter, maintaining better alignment between these two components. Two of the teachers (Teacher D and Teacher M) made changes to their FA design, shifting from tool-oriented to correctness assessing, but mostly maintaining conceptual mastery as their central goal. These teachers, however, demonstrated a shift toward adopting a more CT perspective toward the end of the PD. Two teachers in this group (Teacher G and Teacher J) manifested an inconsistent approach to FA task construction, submitting

https://doi.org/10.1021/acs.jchemed.1c00356
J. Chem. Educ. 2021, 98, 2799−2808

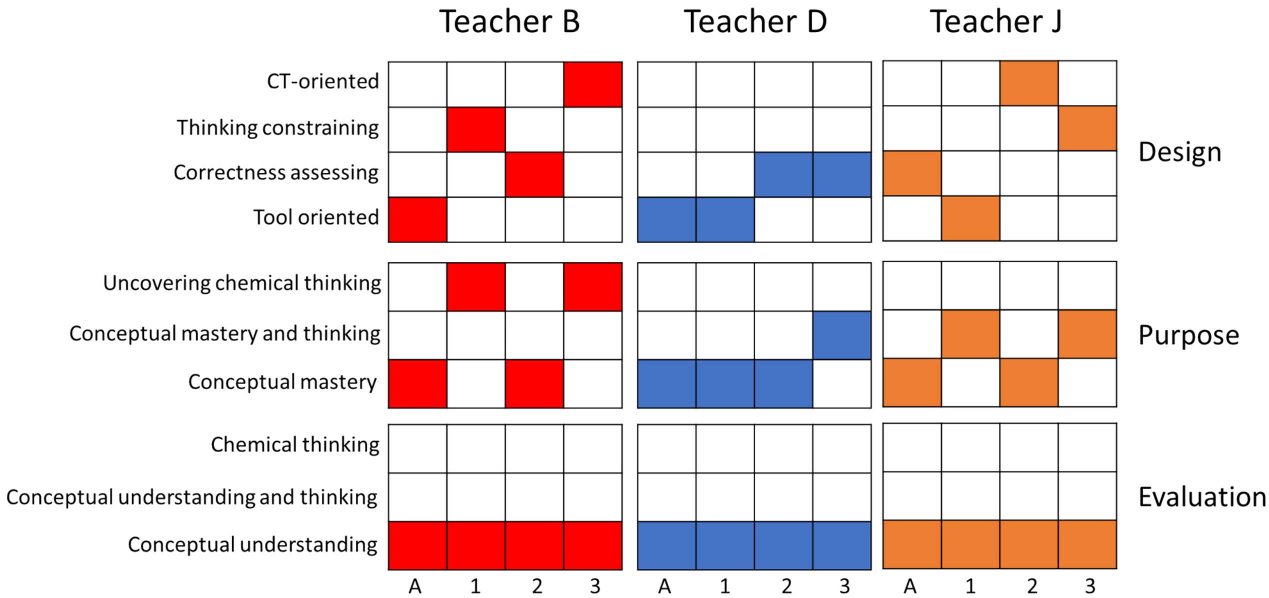


Figure 2. Comparison of a teacher whose design and purpose change in sync, Teacher B; a teacher whose change in purpose is delayed, Teacher D; and a teacher whose changes in purpose change in opposition to her design changes, Teacher J. Note: A = application portfolio chapter, 1 = portfolio chapter 1, 2 = portfolio chapter 2, 3 = portfolio chapter 3. Each of these teachers is representative of a larger group.

activities the stated purpose of which changed in opposition to their task design (e.g., more CT-oriented in design but more focused on conceptual mastery in purpose).

Change in All Areas

Nine participating teachers submitted FA tasks that showed different degrees of change in the three areas under analysis: design, purpose, and evaluation (Figure 3). Five of these teachers submitted at least one FA task with a CT-oriented design at some point during the PD year. In four of these cases, the task submitted in the last portfolio was CT-oriented and designed with the sole purpose of uncovering students' chemical thinking. Three of these teachers also focused their evaluations of students' responses on making sense of student ideas. Four of the nine teachers in this group submitted at least one FA task that reached the level of thinking constraining in its design. Although not fully CT-oriented, the purpose of these tasks varied between conceptual mastery and incorporating some focus on chemical thinking. Similarly, the focus of these teachers' evaluations of student work varied between evaluating conceptual understanding and making sense of student reasoning.

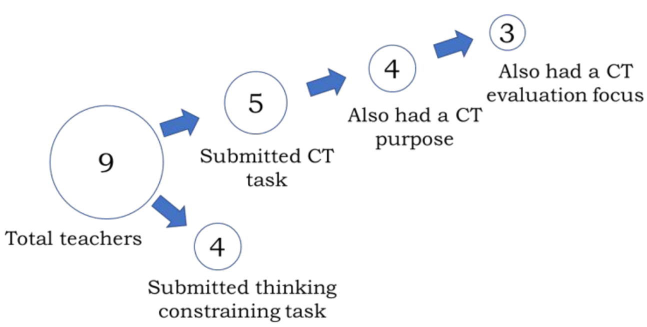


Figure 3. Total number of teachers who made changes in all three areas and the subset of teachers who had a CT focus in one or more areas of analysis.

At the beginning of the PD year, most of the teachers in this category examined students' work with an evaluative stance, looking for what students got right or wrong and paying attention to what the students were not capable of doing or with which work they were struggling. For example, when evaluating the results of a card sort related to energy transfer in his first portfolio chapter, Teacher K wrote "*I don't see evidence of deeper particulate-level thinking*" and "*the student is missing the comparison between energy changes in reactant bonds breaking and product bonds forming.*" In this analysis, the teacher paid attention to what was missing in the students' answers rather than on making sense of students' ideas and looking for opportunities to advance students' chemical thinking. However, the teacher's FA practices changed by the end of the PD program. When evaluating work related to an FA task designed to target students' structure−property relationship thinking, Teacher K wrote in their final FA portfolio chapter:

> *"The student identifies the structures of the two molecules as responsible for more effective bond breaking in the capsaicin molecule. . . .it is tempting to correct the student's apparent misconception about intra and inter molecular forces. A more productive move might be to follow up about why the student believes that similarity in structures will facilitate more effective bond breaking."*

In this example, the teacher demonstrated awareness of their inclination to take an evaluative stance in the analysis of this student's answer but instead suggested a productive way of advancing student thinking by further eliciting and exploring their reasoning. The teacher recognized that the student is trying to reason about the structure−property relationship and that further exploration of this line of thought is more productive than simply telling the student that their answer is not correct. Similar shifts from focusing on correcting student work to uncovering, making sense of, and challenging student thinking were seen in the evaluations of all the teachers in this group at some point throughout the year.

While there were similarities among the shifts demonstrated by teachers in this group, their FA tasks also changed in varied

https://doi.org/10.1021/acs.jchemed.1c00356
J. Chem. Educ. 2021, 98, 2799−2808

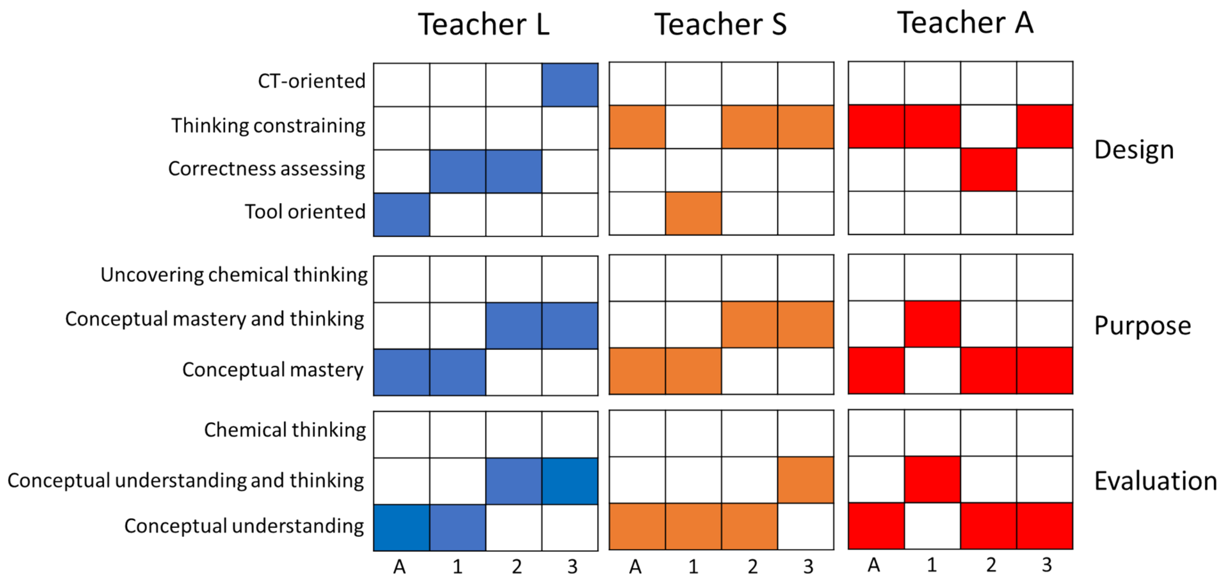


Figure 4. Teacher L displays staggered shifts in his FA practice, first with a change in his design followed by a change in his purpose and evaluation focus. Teacher S shows simultaneous changes in their design and purpose with a delayed change in their evaluation approach. Teacher A was consistent in their task design but alternated back and forth for their purpose and evaluation approach. Note: each of these teachers is representative of a larger group.

ways over the PD year. Two of the teachers (Teacher L and Teacher Q) designed tasks that became more CT-oriented with time, but changes in their purpose and evaluation approach lagged with respect to task design. This discrepancy is illustrated in Figure 4, where we can see how the FAs from Teacher L advance in task design before advancement is seen in the areas of purpose and evaluation. Other teachers, such as Teacher S, tended to create tasks that were aligned in design and purpose, but there was a delay in progress in the evaluation of student work. Some teachers, such as Teacher A, more consistently submitted FA tasks that were CT-oriented, but their purpose and evaluation oscillated between focusing on conceptual mastery and paying attention to students' chemical thinking.

■ DISCUSSION

The central goal of our study was to investigate changes in chemistry teachers' FA practices in three areas (design, purpose, and evaluation of students' work) as they participated in a yearlong PD program guided by the CT framework. All teachers implemented changes in at least one of these areas of analysis, with about half of them creating tasks that demonstrated changes in all three dimensions. Changes in FA design were the most prevalent among participating teachers, which suggests that teachers can more easily incorporate CT-oriented design elements into FA tasks than make changes to their beliefs about the goals of FA and their focus while evaluating student work. This difference may be explained by the fact that changing task design does not necessarily require a shift in teachers' beliefs about the purpose of teaching chemistry, as a CT-oriented FA can still be used to differentiate students based on their conceptual understanding and to elicit their level of content mastery. It is also possible that some of these changes occurred because teachers had more direct access to resources that provided ideas for changes in task design, including CT-oriented FAs designed by the PD leadership team.

Changes in teachers' stated purpose and evaluation approach likely stemmed from teachers beginning to change their perception of the nature of chemistry. Specifically, these teachers appeared to be shifting their views of chemistry as a set of facts to chemistry as a way of thinking about phenomena and of addressing real-world problems. As reported by other scholars,[8,11] most teachers participating in our study came to the PD with approaches to setting FA goals that were narrowly focused on the evaluation of conceptual mastery. Nevertheless, most of them (n = 14) ended the PD year paying some level of attention to student reasoning in chemistry. For example, Teacher B changed her purpose from teaching specific content (i.e., atomic structure) to focusing on students' thinking. This shift required a change in how the teacher viewed chemistry, in this case moving from atomic structure as a set of models to be learned to chemistry as a powerful tool to think about broader questions (i.e., how to identify and differentiate chemical species by using different models to infer and interpret various properties). This change was not easy, as many teachers did not manifest this shift in a consistent manner and the change in purpose was often delayed compared to the change in task design. Several teachers in our study submitted FA tasks that moved back and forth between focusing solely on content mastery and creating some room for exploring chemical thinking. This suggests that a belief that the main goal of chemistry FA is to evaluate whether students know a fact, understand a specific idea, or can demonstrate a particular skill was likely strongly ingrained in these teachers' minds.

Changes in the area of evaluation of student work were the least common among our participating teachers, potentially because such changes require a deep understanding of chemistry and a commitment to a view of chemistry as a way of thinking. Fewer than half of the participants manifested changes in this dimension, and only three teachers presented a single example of FA in which the sole focus was on making sense of student ideas and identifying productive elements to advance student learning. Teacher K acknowledges a common approach to evaluation by stating that it is "tempting" to simply

https://doi.org/10.1021/acs.jchemed.1c00356
J. Chem. Educ. 2021, 98, 2799−2808

correct a student on where they went wrong. This approach, which falls on the conceptual focus end of the spectrum, does not require in-depth understanding of chemistry as a discipline. However, looking for and recognizing productive lines of reasoning (i.e., structure−property relationships), even when the student provides an incorrect answer, demonstrates Teacher K's understanding of the broader goals of learning chemistry and chemical thinking. Most teachers struggled to adopt a less evaluative stance and thus commonly focused on what the students did not get right, what they were not able to do, or what was missing in their work. Few participating teachers, and in limited examples, used the analysis of student work to speculate about what student responses might reveal about student thinking and to reflect on actions that could support student learning. Adopting a more responsive approach in the assessment of student understanding has been shown to be difficult for many teachers,[35] and there are few examples and publicly available resources to guide and support chemistry teachers' development/growth in this area.[20,36] Our project has developed such resources, which are available through the ACS's ChemEd Xchange at https://www.chemedx.org/ACCT. Changes in both purpose and evaluation focus seem to require the teacher to first make changes to their beliefs about the purpose of teaching and learning chemistry. Changes in their beliefs then manifest themselves in the teachers' FA practices.

When changes in FA task purpose and/or evaluation were observed, they tended to mirror changes in task design, in moving either toward or away from a more CT-oriented perspective. This alignment was more common for tasks submitted in the second half of the PD. Nevertheless, changes in design often preceded changes in the other two dimensions. This trend reflects research on the professional development of teachers which indicates that major changes in practice often require over three years of sustained and reflective training.[37] Thus, although teachers' progress was frequently uneven and somewhat inconsistent, it is encouraging to see evidence of advancement in some areas of CT-oriented FA in the relatively short time of the PD under investigation.

■ LIMITATIONS

Although the progress demonstrated by teachers participating in this study is encouraging, we also recognize that there are some limitations to this study. While the sample of participating teachers was fairly diverse, it was somewhat small and thus cannot be used to make generalizable claims. Individual teacher characteristics and contextual factors are likely to affect how teachers change their practices in response to professional development. Data collection occurred at only four points during the PD and may not have been fully representative of teachers' approaches to the design of FA tasks and evaluation of student work during that period. Additionally, these data were collected while teachers were still trying to understand the ideas discussed in their training. Thus, collected FA tasks may correspond to a period in which teachers' understanding was fragmentary, and more time would be required for them to meaningfully and productively integrate new ideas into their FA practices.

■ CONCLUSIONS AND IMPLICATIONS

The major findings of our study suggest that changing chemistry teachers' beliefs and practices in the area of chemistry FA may be challenging. This change may be more easily promoted in some areas, such as FA task design, but additional time and effort is likely to be required to transform teachers' beliefs about the goals of chemistry FA in the classroom and their approaches to the analysis of associated student work. Most teachers in our study did not progress consistently in their adoption of a CT perspective during the PD year. Many of them moved back and forth between focusing on assessing conceptual mastery and paying some attention to student reasoning. This variation may have been due to teachers' incomplete or fragmentary understanding of and commitment to the ideas introduced in the PD program, but it could also have been influenced by the specific chemistry content targeted by the FA tasks that the teachers developed. For example, teachers may find it easier to develop FA tasks with a CT perspective when teaching about structure−property relationships than when teaching stoichiometry. More research is needed to understand how the nature of the content that is taught and the approaches that a teacher traditionally follows in teaching such content interact with teachers' views about the types of FA that are needed to support and advance students' chemical thinking. As is the case for student understanding in chemistry, teacher learning can be expected to progress through nonlinear trajectories and be highly situated and dependent on contextual factors. Multiple and long-term opportunities to practice, receive critical feedback, and reflect in and about action in different contexts are likely needed for teachers to consistently use FA in more responsive and CT-oriented manners.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available at https://pubs.acs.org/doi/10.1021/acs.jchemed.1c00356.

FA portfolio chapter template (PDF)
FA portfolio chapter example (PDF)
Graphs of teacher FA changes (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Timothy N. Abell − *Department of Chemistry, University of Massachusetts Boston, Boston, Massachusetts 02125, United States*; orcid.org/0000-0002-3600-6039; Email: timothy.abell@umb.edu

Author

Hannah Sevian − *Department of Chemistry, University of Massachusetts Boston, Boston, Massachusetts 02125, United States*; orcid.org/0000-0001-7378-6786

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jchemed.1c00356

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful to Vicente Talanquer for his invaluable feedback while writing this manuscript and to Benny Berger for their help in the editing process. This material is based on work supported by the National Science Foundation award DRL-1621228. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

■ REFERENCES

(1) NGSS Lead States. *Next Generation Science Standards: For States, By States*; The National Academies Press: Washington D.C., 2013. DOI: 10.17226/18290.

(2) National Research Council. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*; National Research Council: Washington D.C., 2012. DOI: 10.17226/13165.

(3) Dini, V.; Sevian, H.; Caushi, K.; OrduñaPicón, R. Characterizing the Formative Assessment Enactment of Experienced Science Teachers. *Sci. Educ.* **2020**, *104* (2), 290−325.

(4) Bell, B.; Cowie, B. The Characteristics of Formative Assessment in Science Education. *Sci. Educ.* **2001**, *85* (5), 536−553.

(5) Black, P.; Wiliam, D. Assessment and Classroom Learning. *Assess. Educ. Princ. Policy Pract.* **1998**, *5* (1), 7−74.

(6) Ruiz-Primo, M. A.; Furtak, E. M. Exploring Teachers' Informal Formative Assessment Practices and Students' Understanding in the Context of Science Inquiry. *J. Res. Sci. Teach.* **2007**, *44* (1), 57−84.

(7) Furtak, E. M.; Ruiz-Primo, M. A. Making Students' Thinking Explicit in Writing and Discussion: An Analysis of Formative Assessment Prompts. *Sci. Educ.* **2008**, *92* (5), 798−824.

(8) Harshman, J.; Yezierski, E. Guiding Teaching with Assessments: High School Chemistry Teachers' Use of Data-Driven Inquiry. *Chem. Educ. Res. Pract.* **2015**, *16* (1), 93−103.

(9) Kloser, M.; Borko, H.; Martinez, J. F.; Stecher, B.; Luskin, R. Evidence of Middle School Science Assessment Practice from Classroom-Based Portfolios. *Sci. Educ.* **2017**, *101* (2), 209−231.

(10) Kang, H.; Windschitl, M.; Stroupe, D.; Thompson, J. Designing, Launching, and Implementing High Quality Learning Opportunities for Students That Advance Scientific Thinking. *J. Res. Sci. Teach.* **2016**, *53* (9), 1316−1340.

(11) Sandlin, B.; Harshman, J.; Yezierski, E. Formative Assessment in High School Chemistry Teaching: Investigating the Alignment of Teachers' Goals with Their Items. *J. Chem. Educ.* **2015**, *92* (10), 1619−1625.

(12) Harshman, J.; Yezierski, E. Characterizing High School Chemistry Teachers' Use of Assessment Data: Via Latent Class Analysis. *Chem. Educ. Res. Pract.* **2016**, *17* (2), 296−308.

(13) Abell, T. N.; Sevian, H. Analyzing Chemistry Teachers' Formative Assessment Practices Using Formative Assessment Portfolio Chapters. *J. Chem. Educ.* **2020**, *97* (12), 4255−4267.

(14) Murray, S. A.; Huie, R.; Lewis, R.; Balicki, S.; Clinchot, M.; Banks, G.; Talanquer, V.; Sevian, H. Teachers' Noticing, Interpreting, and Acting on Students' Chemical Ideas in Written Work. *J. Chem. Educ.* **2020**, *97* (10), 3478−3489.

(15) Talanquer, V.; Bolger, M.; Tomanek, D. Exploring Prospective Teachers' Assessment Practices: Noticing and Interpreting Student Understanding in the Assessment of Written Work. *J. Res. Sci. Teach.* **2015**, *52* (5), 585−609.

(16) Russ, R. S.; Coffey, J. E.; Hammer, D.; Hutchison, P. Making Classroom Assessment More Accountable to Scientific Reasoning: A Case for Attending to Mechanistic Thinking. *Sci. Educ.* **2009**, *93* (5), 875−891.

(17) Coffey, J. E.; Hammer, D.; Levin, D. M.; Grant, T. The Missing Disciplinary Substance of Formative Assessment. *J. Res. Sci. Teach.* **2011**, *48* (10), 1109−1136.

(18) Hattie, J.; Timperley, H. The Power of Feedback. *Rev. Educ. Res.* **2007**, *77* (1), 81−112.

(19) Otero, V. K.; Nathan, M. J. Preservice Elementary Teachers' Views of Their Students' Prior Knowledge of Science. *J. Res. Sci. Teach.* **2008**, *45* (4), 497−523.

(20) Abell, T. N.; Sevian, H. Analyzing Chemistry Teachers' Formative Assessment Practices Using Formative Assessment Portfolio Chapters. *J. Chem. Educ.* **2020**, *97* (12), 4255−4267.

(21) Sevian, H.; Talanquer, V. Rethinking Chemistry: A Learning Progression on Chemical Thinking. *Chem. Educ. Res. Pract.* **2014**, *15* (1), 10−23.

(22) Ngai, C.; Sevian, H.; Talanquer, V. What Is This Substance? What Makes It Different? Mapping Progression in Students' Assumptions about Chemical Identity. *Int. J. Sci. Educ.* **2014**, *36* (14), 2438−2461.

(23) Ngai, C.; Sevian, H. Capturing Chemical Identity Thinking. *J. Chem. Educ.* **2017**, *94* (2), 137−148.

(24) Stains, M.; Escriu-Sune, M.; De Santizo, M. L. M. A.; Sevian, H. Assessing Secondary and College Students' Implicit Assumptions about the Particulate Nature of Matter: Development and Validation of the Structure and Motion of Matter Survey. *J. Chem. Educ.* **2011**, *88* (10), 1359−1365.

(25) Talanquer, V. Progressions in Reasoning about Structure-Property Relationships. *Chem. Educ. Res. Pract.* **2018**, *19* (4), 998−1009.

(26) Weinrich, M. L.; Talanquer, V. Mapping Students' Conceptual Modes When Thinking about Chemical Reactions Used to Make a Desired Product. *Chem. Educ. Res. Pract.* **2015**, *16* (3), 561−577.

(27) Yan, F.; Talanquer, V. Students' Ideas About How and Why Chemical Reactions Happen: Mapping the Conceptual Landscape. *Int. J. Sci. Educ.* **2015**, *37* (18), 3066−3092.

(28) Banks, G.; Clinchot, M.; Cullipher, S.; Huie, R.; Lambertz, J.; Lewis, R.; Ngai, C.; Sevian, H.; Szteinberg, G.; Talanquer, V.; et al. Uncovering Chemical Thinking in Students' Decision Making: A Fuel-Choice Scenario. *J. Chem. Educ.* **2015**, *92* (10), 1610−1618.

(29) Cullipher, S.; Sevian, H.; Talanquer, V. Reasoning about Benefits, Costs, and Risks of Chemical Substances: Mapping Different Levels of Sophistication. *Chem. Educ. Res. Pract.* **2015**, *16* (2), 377−392.

(30) Ambitious Science Teaching. Eliciting Students' Ideas. https://ambitiousscienceteaching.org/eliciting-students-ideas-2/ (accessed 2020-10-01).

(31) Martínez, J. F.; Borko, H.; Stecher, B.; Luskin, R.; Kloser, M. Measuring Classroom Assessment Practice Using Instructional Artifacts: A Validation Study of the QAS Notebook. *Educ. Assess.* **2012**, *17* (2−3), 107−131.

(32) Borko, H.; Stecher, B. M.; Alonzo, A. C.; Moncure, S.; Mcclam, S. Artifact Packages for Characterizing Classroom Practice: A Pilot Study. *Educ. Assess.* **2005**, *10* (2), 73−104.

(33) Martínez, J. F.; Borko, H.; Stecher, B. M. Measuring Instructional Practice in Science Using Classroom Artifacts: Lessons Learned from Two Validation Studies. *J. Res. Sci. Teach.* **2012**, *49* (1), 38−67.

(34) Kyngäs, H.; Kääriäinen, M.; Elo, S. The Trustworthiness of Content Analysis. In *The Application of Content Analysis in Nursing Science Research*; Springer International Publishing: Cham, 2020; pp 41−48. DOI: 10.1007/978-3-030-30199-6_5.

(35) Dini, V.; Sevian, H.; Caushi, K.; OrduñaPicón, R. Characterizing the Formative Assessment Enactment of Experienced Science Teachers. *Sci. Educ.* **2020**, *104* (2), 290−325.

(36) Assessing for Change in Chemical Thinking. www.chemedx.org/ACCT (accessed 2020-10-01).

(37) Banilower, E. R.; Smith, S. P.; Weiss, I. R.; Malzahn, K. A.; Campbell, K. M.; Weis, A. M. *Report of the 2012 National Survey of Science and Mathematics Education*; Horizon Research, Inc.; Chapel Hill, NC, 2013.