JOURNAL OF CHEMICAL EDUCATION

pubs.acs.org/jchemeduc Article

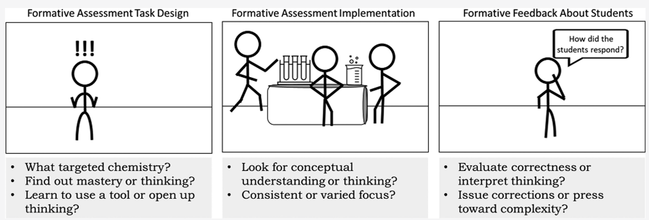
Analyzing Chemistry Teachers' Formative Assessment Practices Using Formative Assessment Portfolio Chapters

Timothy N. Abell and Hannah Sevian*

Cite This: J. Chem. Educ. 2020, 97, 4255−4267 Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

ABSTRACT: The effective use of formative assessment (FA) has been demonstrated to confer positive impacts on student learning. To understand why and how FA works, it is necessary to characterize teachers' FA practices, but because both teaching practice and learning depend on the nature of the discipline, there are disciplinary aspects to examining this. This study aimed to develop an analysis of chemistry teachers' FA practices through the lens of the chemical thinking framework. Two cohorts of middle and high school science teachers participated in year-long professional development with the goal of improving their FA practices in teaching chemistry. Each teacher submitted FA portfolio chapters throughout the year. To develop an approach to use in ongoing research that will analyze teachers' progress across the year, the final FA portfolio chapters of participants ($N = 13$) were analyzed to characterize FA task design, the teacher's purpose in implementing the FA, and how the teacher evaluated student work. FA tasks were found to range from revealing students' mastery of concepts to uncovering students' chemical thinking. Teachers also demonstrated a range of purposes behind their use of the FA tasks, and a range of focuses when evaluating student work. Correspondence among the FA task, a teacher's purpose for its use, and the teacher's evaluation approach revealed patterns that echo the broader research in science education, but with instantiation in chemistry. Ways for teachers to assess and diversify their own FA practices based on these findings are presented.



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

FEATURE: Chemical Education Research

■ INTRODUCTION

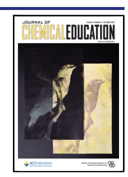
Formative assessment (FA) is an important tool for teachers to support student learning. Bell and Cowie[1] define FA as

> *the process used by teachers and students to recognize and respond to student learning in order to enhance that learning, during the learning.*

This definition of FA is broad, including activities that are planned ahead of time by the teacher and then carried out with students, discussions that occur spontaneously as activities occur in the classroom, and review and feedback by teachers on students' written work. For example, teachers often provide immediate feedback during an FA activity, and offer written, delayed, feedback after the FA through the evaluation of the written work.[2] FA can also happen at any time during a lesson when a teacher notices ideas and thinking shared by students and then responds to the students in real-time, offering immediate feedback.[1,2] FA has been shown to have a significant impact on student learning.[3,4] While science teachers' FA practices are widely studied in science education, both teaching and learning have discipline-specific aspects, so it is also important to consider the FA practices of teachers through a disciplinary lens.

In chemistry, researchers have investigated students' perceptions of different types of FA,[5] characterized college teachers' familiarity with assessment terminology,[6,7] and presented a framework for writing three-dimensional learning assessment questions.[8] Few chemistry-specific studies have analyzed teachers' FA practices. Harshman and Yezierski analyzed teachers' assessment practices by looking at relationships between teachers' stated goals and the items they selected.[9−11] They found that teachers tend to set larger unit goals and not goals for individual assessments, but the goals they do set tend to focus on conceptual understanding.[9,11] To evaluate student work, teachers used descriptive statistics (percentage correct) to judge their students' level of understanding. Overall, there was rather poor alignment between what the teachers wanted to assess and what the items they chose actually assessed.[9,10] The authors did not

Received: April 16, 2020
Revised: October 7, 2020
Published: November 16, 2020



characterize the teachers' FA practices beyond the alignment between goals and items. Furthermore, the authors were only able to gather rather ambiguous data through their interviews, and they called for further inquiry into chemistry teachers' assessment practices within a disciplinary context using a data collection method that would allow teachers to elaborate on their assessment practices.

A major approach to examining FA practice in science education is to use unit portfolios or student notebooks.[12−16] Typically, a portfolio is a collection of artifacts of FA that includes instructional and planning materials, assessments, examples of student work, a teacher's feedback and comments on student work, and the teacher's reflections on the lesson. Researchers have also complemented the analysis of artifacts with classroom observations to include data related to the enactment of the FA. While portfolios of entire units have been found to be a viable way to understand a teacher's practices, this method places a great deal of time burden on teachers to collect and assemble full unit portfolios. Borko et al.[13] made recommendations for ways to reduce the number of artifacts and days on which artifacts are collected.

When analyzing FA practice, focus has concentrated on two main areas: FA design and teachers' evaluation of student work. One of the most important aspects of designed FA is the task itself, as it affects the entire FA process, including the insight gained into students' thinking.[2,9,12,17] Kang et al.[17] found that the cap for the intellectual demand for an FA is set by the design of the task; i.e., the intellectual demand can remain at the level of the task, but it much more typically decreases across enactment of the FA, from the launch through implementation of the FA and then evaluation of student work. In chemistry, task design was shown to limit the conclusions the teacher was able to draw about their students because the items used did not assess the teachers' stated goals.[9] The design of FA tasks has been shown to correlate with overall practices of teachers. Teachers who designed tasks that asked students to make connections between concepts and to generate explanations tended to have better overall assessment practices.[12] Two major findings about FA practice transcend many studies.[2,3] To uncover students' thinking in writing, FA prompts should be both open-ended and familiar to students. After FA prompts are answered, a teacher must provide some feedback on student work so that students can act upon it.

Previous research has examined how teachers notice and interpret student work and what kinds of feedback they provide to students. When noticing student work, novice science teachers tend to describe what students write, often simply rewording students' responses.[18] More advanced approaches, while rare in novice teachers, involve making inferences about student understanding, in particular, by identifying underlying assumptions.[18] When interpreting students' work, novice teachers tend to focus on the "correctness" of the students' answers, rather than uncover the productive ways of thinking that students demonstrate.[18−22] This may be, in part, due to a lack of clarity on the part of teachers between eliciting students' thinking and asking for correct answers.[23] Talanquer et al.[18] found that some advanced prospective science teachers were able to make sense of students' thinking when evaluating their work, but they tended to be inconsistent in this focus. These studies converge on the importance of shifting teachers' focus toward noticing students' underlying thinking because this enables teachers to leverage productive aspects of the students' own reasoning.[18,19]

The planning of FA and evaluation of student work have domain-general and domain-specific aspects. To fill the data collection problem posed by the domain-specific research conducted in chemistry about teachers' assessment practices, this study draws on science education methods, such as portfolios. Research that has examined portfolios, as well as most studies of teachers' evaluation of student work, has been carried out through domain-neutral lenses, focusing on science in general or considering data collected across teachers from many areas of science. The goal of this study is to develop a domain-specific analysis of formal FA practices to uncover what chemistry teachers emphasize as important throughout the FA process, from FA design to evaluation of student work. We envision two audiences for this work: (1) teachers who want to focus their FA practice more on students' chemical thinking and (2) researchers who are interested in investigating chemistry-related FA task design and teachers' evaluation of student work in chemistry.

■ THEORETICAL FRAMEWORK

This research is guided by the chemical thinking framework.[24] In order to promote students' abilities to make informed decisions, build justifications, evaluate outcomes, and test ideas in the context of personal, societal, and global concerns, a change of focus from teaching chemistry as a body of knowledge to teaching chemistry as a way of thinking is needed. This framework aims to shift the focus of the chemistry curriculum from being a collection of topics that need to be learned toward being a set of practices of chemistry, in line with current policy.[25] The framework is centered on core questions that chemistry is uniquely positioned to address. The six core questions of chemical thinking are the following:

1. How do we identify chemical substances? (chemical identity)[26,27]
2. How do we predict properties of materials? (structure−property relationships)[28,29]
3. Why do chemical processes occur? (chemical causality)[30,31]
4. How do chemical processes occur? (chemical mechanism)[30,31]
5. How can we control chemical processes? (chemical control)[32]
6. How do we evaluate the consequences of chemically transforming matter? (benefits−costs−risks)[33,34]

Chemical thinking is developed through the application of chemistry knowledge in order to analyze, synthesize, or transform matter as a chemist would. Because all of the practices of chemistry, as instantiated in the six core questions above, involve the application of knowledge related to many topics (e.g., nomenclature, stoichiometry calculations, kinetic molecular theory applications, reaction rate laws, electrophiles and nucleophiles), they are also disciplinary concepts that cut across all of chemistry. When teachers attend to how students think about these crosscutting disciplinary concepts, they can identify productive resources that students have for practicing chemistry (further explicated in the references associated with each chemical thinking question above), while they also attend to supporting students in developing competence in the use of facts and skills (chemistry knowledge) that are applied in

https://dx.doi.org/10.1021/acs.jchemed.0c00361
J. Chem. Educ. 2020, 97, 4255−4267

Table 1. Demographic Information about Participating Teachers

Teacher	Years Teaching	Primary Subject	Grade Level	Other Subjects Taught	Type of School
A	6	Chemistry	10, 11, 12	Physics	Large comprehensive high school
B	5	Chemistry	10	None	English learners high school
C	6	Chemistry	10, 11, 12	Biology, Anatomy & Physiology	Inclusion high school
D	11	Chemistry	11, 12	Earth & Environmental Science	Large selective high school
E	3	Chemistry	11, 12	Biology, Earth & Environmental Science	Small learning community high school
G	12	General Science	6, 7, 8	None	Spanish bilingual middle school
M	7	Chemistry	10, 11	Biology, Biotechnology	Large selective high school
P	3	Biology, AP Biology	11, 12	Chemistry	Vocational−technical high school
Q	6	Chemistry	9, 10, 11, 12	None	Arts magnet high school
R	7	Chemistry, AP Chemistry	10, 11, 12	Earth & Environmental Science	Large selective high school
S	3	Chemistry	11, 12	None	Small learning community high school
T	14	Chemistry, AP Chemistry	11, 12	Physics	Small learning community high school
U	5	General Science	6, 7, 8	Biology, Physics, Earth & Environmental Science, Technology/Engineering	Spanish bilingual middle school

answering core questions of chemistry. In this way, teachers leverage these productive resources toward developing students' chemical thinking. Using the chemical thinking framework allows for the analysis of teachers' FA practices through a domain-specific lens that prioritizes the development of students' chemical thinking and reasoning over the evaluation of right and wrong answers, which is often employed when teaching with a topic-focused approach.

■ RESEARCH QUESTIONS

The goal of this research was to develop a method to analyze what teachers emphasize in their planning and reflecting on FAs. Our motivation in developing this analysis method for FA portfolios is to be able to analyze teachers' progress across a year-long professional development program to look for changes in teachers' emphasis toward a chemical thinking approach. The characterization of FA portfolios is also more broadly useful for two reasons: (1) it offers a mechanism for collecting data on FA practice which includes necessary information for such analysis, and (2) it provides teachers with a way to analyze their own FA designs and evaluation of student work to gauge the extent to which they are emphasizing chemical thinking.

The study was guided by the following questions:

1. How can what chemistry teachers emphasize in planning and reflecting on FA be characterized?
2. What types of emphases can be illuminated using this characterization?

■ PARTICIPANTS

The teachers in this study ($N = 13$) were participants in a year-long professional development (PD) program. Teachers were selected for the program on the basis of a double-blind review of applications by teachers and researchers in the project leadership. The review prioritized school variety, diversity of teachers' situations, and range of years of experience in selecting each cohort. Review considerations also included the potential for growth in diversifying teachers' assessment stances (i.e., the extent to which teachers focused on judging correctness or on students' sense-making[18]) that were evident in two short essays and an example FA with the teacher's reflections on one student's work. Eligibility to apply included being a high school chemistry or middle school science teacher in a specific large urban New England public school system, having at least three years of teaching experience, and obtaining approval of the principal. Artifacts from two cohorts, in 2017−2018 ($n = 5$) and 2018−2019 ($n = 8$), are included in the sample of data analyzed. The 13 teachers in this study were either high school chemistry ($n = 10$), high school biology ($n = 1$), or middle school science ($n = 2$) teachers. (The biology teacher applied and was selected while being a chemistry teacher, and then was reassigned to teach biology after the cohort year began; this teacher collected all data during chemistry-related lessons in biology.) The teachers had between 3 and 14 years of teaching experience, with an average of 6.8 years of experience, taught in different types of schools, and also taught other subjects (Table 1). Demographic information was collected to show the diversity among participants. The sample size is too small to be used to draw conclusions based on demographics. The study was approved by the university's IRB and the school district's research compliance office. All teachers provided written consent to participate in the study. The teachers were compensated either by a stipend or graduate-level credits in chemistry. The students provided written consent if over 18; students under 18 provided assent, and consent was provided by a parent or guardian. To preserve confidentiality, teachers are referred to by single letters that are not related to their actual names.

The goals of the PD were to strengthen and diversify teachers' formative assessment practices in three areas: selection and design of formative assessment tools, evaluation of students' responses, and decision-making based on what is learned from interactions with and responses of students during FA. Regarding these, the PD supports teachers shifting from using FAs designed exclusively to differentiate students on the basis of ability level, toward FAs that also can elicit and promote students' chemical thinking. The PD also aims to shift teachers' focus from evaluating student work as correct or incorrect, toward increased interpreting of students' underlying thinking and identifying students' own productive resources. Third, the PD intends to focus teachers on using what they learn from the evaluation of their students' work in order to build upon students' productive ideas rather than only relecturing the main ideas. This research centers on characterizing teachers' FA practices in relation to the first two areas.

METHODS

Each teacher assembled an FA portfolio consisting of four chapters spaced across the year, beginning with one submitted with their application and concluding with one near the end of the school year. The FA portfolio design was adapted from Scoop Notebooks,[14,35] which have been used in measuring instructional practices of science and mathematics teachers. For each portfolio chapter, teachers were asked to provide a blank copy of the FA task, with explanations of their purpose in using the FA, where the FA fit into the curriculum, and expectations for a proficient student's answers and confusing aspects that students may encounter. The teachers also provided deidentified work from three consenting students that was representative of the range of work by the class, along with how the teacher evaluated these students' work. As a part of their evaluation, some teachers provided hypothetical feedback about what they may have said or asked a student based on their answers. In almost all cases, this feedback was never actually presented to the students and was rather part of the teachers' reflections on students' work. Finally, teachers noted how they proceeded with the lesson immediately after the FA or in the next class, and they reflected on what worked well and what they would change. The template version of the FA portfolio chapter and an example that was provided to teachers can be found in the Supporting Information. To develop a method of characterizing FA portfolio chapters, we analyzed the final FA portfolio chapters submitted by each participant because these were expected to represent the greatest range of assessment emphases. These were analyzed in three parts: (1) the design of the FA, (2) the teacher's intentions for using the FA, and (3) the teacher's evaluation of and feedback on student work. This analysis only focuses on the delayed written evaluation and feedback provided by the teacher in an FA portfolio chapter. The in-the-moment feedback that is also part of FA was captured with classroom video observation and is the subject of a separate study.

The FA design was analyzed following two criteria from Ambitious Science Teaching[36] which were incorporated into the PD: accessibility and power to reveal students' thinking. Accessibility is the degree to which students have enough knowledge about the questions or task to be able to engage in the work. The FA's accessibility was evaluated by considering not only the questions asked but also the context in which the questions were embedded and the FA's place in the curriculum with respect to the students' prior learning experiences in school. Each FA was classified as having "High" or "Low" accessibility. The FA's power to reveal students' thinking was considered with respect to the chemical thinking strand that most closely related to the ideas and reasoning students were asked to engage with in the FA.[24] To gauge this, we examined the task and questions presented to the students, along with the types of responses that students provided. The task and questions were evaluated to determine the extent to which the FA had the potential to reveal students' chemical thinking. The students' responses were then used to determine to what extent this potential was met. FA tasks were grouped on the basis of similarities in their accessibility and power to reveal students' chemical thinking.

A characterization of a teacher's purposes for using an FA was synthesized from several questions in the FA portfolio chapter. These questions asked the teacher 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. From these, a narrative was written for each teacher describing the teacher's goal(s) for using this specific FA. These narratives were then grouped on the basis of commonalities of goals.

The teacher's evaluation of student work was characterized by analyzing how and what the teacher noticed and interpreted on the basis of the student responses in the FA portfolio chapter. For both metrics, noticing and interpreting, each teacher was characterized as "Minimal" or "Extensive". If a teacher demonstrated characteristics at both extremes, the teacher was considered to be between the two extremes. When *noticing* a student's work, "Minimal" was recorded if the teacher only described what a student had written by either directly quoting the student or rewording what the student had written.[18] In such cases, the teacher did not go beyond surface level in examining the student work. An "Extensive" code was recorded when a teacher described inferences about the student's understanding.[18] To do this, a teacher speculated on assumptions relied upon by a student or identified ideas the student appeared to be considering in order to develop an answer. When *interpreting* students' thinking, coding ranged from "Corrective" to "Sense-making". Corrective interpreting involved comparing student responses to canonical knowledge to determine whether they were right or wrong, judging a student's level of understanding, or explaining the route a student took to arrive at an incorrect answer (e.g., the student failed to convert Celsius to Kelvin).[18,19,21,37] "Sense-making" interpreting involved an effort to understand students' underlying thinking related to a chemical thinking strand.[18] This went beyond whether a student response was right or wrong; rather than emphasizing judgment of correctness of a student response, the teacher focused on hypothesizing the underlying thinking that may have led the student to reason in this way. This analysis was carried out separately for each of the three student work samples evaluated by a teacher. A narrative was then developed on the basis of the teacher's tendencies when noticing and interpreting the students' work. Similar narratives were grouped, and themes were developed on the basis of focuses in each group in evaluating student work.

In developing the analysis method, we paid attention to important areas of trustworthiness: dependability, credibility, transferability, and confirmability.[38] Dependability was ensured through a detailed description of the data collection and analysis process (detailed above). Peer examination was used to establish the credibility of the results. Debrief sessions were held with a group that included six individuals who were postdocs and graduate students and were not associated with the data collection or the development of the analysis. During the debrief sessions, all attendees coded data using the coding scheme and provided feedback. This helped ensure that the coding scheme did not have blind spots and that similar findings were drawn by several people. The diverse backgrounds of the participants (Table 1) support the transferability of the results and conclusions of this study. After

https://dx.doi.org/10.1021/acs.jchemed.0c00361
J. Chem. Educ. 2020, 97, 4255−4267

Table 2. Synopsis of Tasks in the FA Portfolio Chapters

Teacher	Starting Materials, Scenario, or Activity	Corresponding Formative Assessment Task
A	Students are presented with a photograph of a puddle of water and a particulate representation of water. They are told that the puddle evaporates and are given four particulate representations of gaseous water.	Students are asked to determine whether they agree or disagree with each representation and then to explain why.
B	Students are given photographs of water in solid, liquid, and gas forms.	Students are asked to list the ways in which the three forms of water differ, draw a particulate representation of water in each phase, and list the phases in increasing order of energy.
C	This FA consists of a prelab and lab activity.	In the prelab, students are asked to write and balance a chemical reaction, identify the type of chemical reaction, and determine the limiting reagent.
		The aim of the lab activity is to determine the empirical formula of zinc chloride.
D	Students are given cards representing the components of a container of pure water (H_2O, H^+, OH^-) as well as extra H^+ and OH^- cards to represent an Arrhenius acid and base, respectively.	Students are tasked with manipulating the cards to represent how a solution would become more acidic or basic through addition of the ions.
E	A written scenario describes three teachers of differing sizes who consume the same amount of caffeine and the smallest teacher dies.	The task presents a series of questions about why only one teacher died, what caused the teacher to die, and why some compounds are toxic for some but not others.
G	A drawing is given of a town at the bottom of a mountain with two points labeled (one at the top of the mountain and one in the town).	Students are asked about the relative amounts of phosphorus at each point, what they expect to happen over time to the phosphorus at each point, and their proposed ways to control the runoff of phosphorus.
M	Four different colognes/perfumes, a list of their ingredients, and some of their structures are given to students.	Students are asked to look for patterns among the lists of ingredients, molecular structures, and functional groups to determine what has an impact on the scent.
P	Students are presented with a scenario in which a science fair project is being planned.	Students are given the different variables used in the study and asked to identify and define the independent and dependent variables.
Q	This FA includes two components: a Do Now at the beginning of class, and an Exit Ticket at the end.	Both ask students what they know about acids and what identifies a substance as an acid. During the class activity in between, students engage in an activity that involves an acid but are not taught explicitly about acids.
R	Students conduct a laboratory experiment using titration and a color changing indicator.	When concluding the experiment, students are asked questions about the use of diprotic acids instead of monoprotic and about the effects of adding water to the titrant, and they are asked about the calculation of the moles of acid, the method to fix an overtitrated solution, and the identification of sources of error.
S	Students are given colored powdered drink mix and water.	Students are asked to make an "ideal" solution based on taste. Students are then asked to manipulate their solutions by adding more water or more drink mix powder and then explain the changes that occur through writing and particulate representations.
T	Students are given three solutions of NaCl in various volumes and concentrations.	Students are then asked to determine which solution is most concentrated, which has the most moles of solute, and which has the largest mass of solute.
U	Students are presented with a drawing of an erupting volcano.	Students are asked to draw and explain in words what they think occurred underground that caused the eruption.

https://dx.doi.org/10.1021/acs.jchemed.0c00361
J. Chem. Educ. 2020, 97, 4255−4267

completing and submitting their FA portfolio chapters, a subset of the teachers were asked to provide additional reflections on their FA practices associated with their final FA portfolio chapter. The teachers' own analyses of their focuses during the evaluation of students' work were compared to the results of the coding scheme as a way of establishing confirmability (see Supporting Information).

■ RESULTS AND DISCUSSION

The teachers developed chemistry-based FAs based on a variety of tasks (summarized in Table 2). These ranged across content in middle school science, high school chemistry, chemistry topics in biology, and AP chemistry. They included a wide variety of models (e.g., Arrhenius acid−base, particulate representations of phase changes, reaction types), contexts (walking around the school, the kitchen at home, a chemistry laboratory, a science fair, a rural watershed, a geological formation), and modes of activity (drawing a representation, animating a reaction, analyzing a story, conducting laboratory work, making drinks from colored powdered mix). The FA activities included both individual and group work. Most student work involved written explanations and drawings of representations, but some student work included still photos, e.g., when there were posters, animations, or lab setups produced by students.

Analysis of the teachers' FA portfolio chapters resulted in differences in design, purpose, and evaluation of student work, presented in the following three sections. The categorizations of approaches within these are not intended to denote "good" or "bad" ways of enacting FA. Each design, purpose, and evaluation approach can be used by teachers in valuable ways to achieve different goals. Because a goal of the PD was to diversify teachers' FA practice, our interest in developing this analysis approach was to be able to capture the varied ways in which teachers enact FA. As explained in the Theoretical Framework section, however, the rationale for offering the PD was to increase teachers' emphasis on chemical thinking. It is easier to adapt a chemical thinking task into a content and skills-focused task than it is to do the reverse. In our framing of implications, we comment first on affordances, and then on ways of shifting FA practices from a content/skills focus to chemical thinking.

Formative Assessment Design

The FA tasks (Table 2) were analyzed to determine the extent to which they were accessible and capable of eliciting students' chemical thinking.

Tool-Oriented FA Design. Tasks with a tool-oriented design focus on developing a specific skill or practicing the use of a specific concept. Usually, these involve questions that have one correct answer. Tool-oriented tasks often involve specific prior knowledge and are not easily accessible to a student who has not had learning experiences related to that knowledge. Four teachers in our sample developed or selected tool-oriented tasks. For example, teacher T, a high school teacher, designed an FA around students' understanding of concentration and the ability to apply the formula $C = n/V$ (Figure 1). The task presented students with three solutions of NaCl in various volumes and concentrations. Students were asked to determine which solution is most concentrated, which had the most moles of solute, and which had the largest mass of solute, and to explain why these were the answers for each question. Each of the questions has only one correct answer, and the route to each correct answer is through application of the same equation. Specific prior knowledge is needed by students to be able to engage in this task, including recognition of units of measure (molarity, volume) and ability to report concentration (e.g., 0.6 M NaCl), use of the periodic table to determine formula weight, and knowledge of how to convert moles to grams. This task allows a teacher to assess a student's understanding of a concept, concentration in this case, and/or proper application of an equation. It does not permit discerning how a student thinks about using the concept as a tool to explain a phenomenon, nor what the student considers to be the meaning or purpose of the equation. Other teachers with similar design features in their tasks asked students to identify types of reactions and charges on molecules/ions, or to define terms, such as dependent vs independent variables.

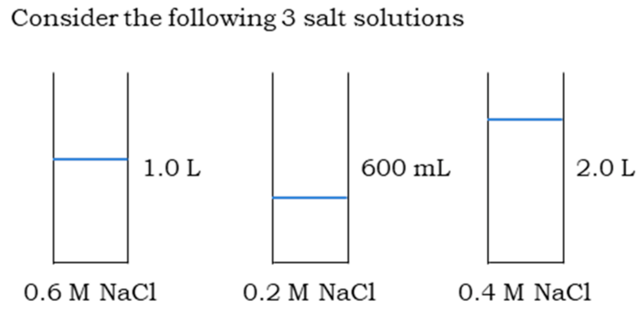


Figure 1. Different concentration solutions of NaCl provided to students on teacher T's FA.

Tool-oriented designs limit the purposes for which the FA can be used to determine whether students recall a specific concept or how students apply an algorithm. This is certainly appropriate at times, as learning chemistry includes understanding concepts and practicing skills. For example, a tool-oriented design is appropriate for practicing drawing Lewis structures. This type of task does not, in itself, offer insight into how a student thinks about chemistry. A teacher might modify a tool-oriented task depending on what questions are asked during the implementation of the FA activity; however, as others have found[17] and we also report later, this rarely occurs. Tool-oriented tasks can also be adapted to focus on chemical thinking, rather than on concepts and skills practice, by embedding the use of the tool in a phenomenon in chemistry that can be explained by the application of the tool. For example, if given a table of concentrations and rates typical in traditional kinetics problems, students could be asked to propose three different ways to control the rate of the reaction, which would involve reasoning using the rate equation that students must first derive from the data.

Correctness-Assessing Structure in FA Design. Tasks of this design are anchored in examining correctness in how students explore and explain a phenomenon. Unlike tool-oriented tasks, which center on reciting knowledge or applying skills, it is the structures of these tasks that constrain them to assessing only correctness. Two teachers submitted FA tasks of this design. Teacher D, a high school chemistry teacher, designed an FA that presented students with cards representing the components of a container of pure water (H_2O, H^+, OH^-), along with extra H^+ and OH^- cards. The students were tasked with manipulating the cards to represent how a solution would become more acidic or basic through the addition of only one kind of extra ion. This task has the potential to explore students' thinking about a mechanism that leads to a more acidic or basic solution. However, the questions in the task asked about how many H^+ and OH^- cards remained after each step, whether the solution was acidic or basic, and what general trend was observed when adding acid or base (e.g., "What happens to the amount of hydrogen ions when you add acid to pure water?"). These questions constrain students to share their thinking only about what the teacher wants students' attention focused on, and the questions can be answered with a single number or word, leaving little room for insight into how a student arrived at an answer. Like the tool-oriented category, specific prior knowledge is necessary for a student to engage in the FA. For example, in teacher D's FA, students needed to know that H^+ and OH^- are representations of Arrhenius acids and bases, and that more of one or the other increases an acid or base solution's strength.

These tasks have the potential to reveal students' chemical thinking; however, either the ways that the questions were asked or the overall design did not provide opportunities for students to do so. Tasks with correctness-assessing structure can be useful when guiding students in a way to think through logical steps in reasoning. Shifting such tasks toward a chemical thinking emphasis requires adding more open-ended questions (e.g., What happens when H^+ ions are added to pure water? How does the solution become more acidic?). Such questions would open opportunities for students to use their conceptual understanding to explain how they envision a mechanism underneath a phenomenon.

Thinking-Constraining FA Design. Tasks with this design constrain the options students can select. Often, these options are common misconceptions. Two teachers' FA tasks presented this design. Teacher A's task asked what happens to water molecules when they evaporate from a puddle. Instead of asking students to draw representations of how they envisioned water molecules before and after evaporation, the task provided students with different representations (Figure 2)

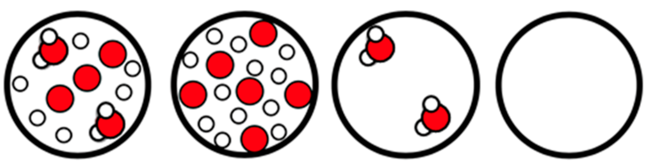


Figure 2. Four representations of water evaporating that students were given in teacher A's FA.

and asked students to judge their correctness and comment on them. While the task is about chemical mechanism (How do the arrangements of water molecules change as they evaporate?), the thinking targeted by the task constrains students to the question of chemical identity (Which representations show water molecules?). Although students may apply chemistry knowledge to decide whether they agree with a representation, the answer options cue students to pay attention to different features than the FA task evokes rather than allowing students to generate representations that better match their thinking related to the question about water evaporating. The forced-choice FA design also constrains students to evaluate known misconceptions rather than share ideas which may not be among the options.

Thinking-constraining task designs can offer a dipstick to gauge whether known misconceptions are present in a class, but like the correctness-assessing structure, the design also limits the ways in which students can respond. In this case, it limits by reducing the avenues available to students in explaining a phenomenon. Unlike correctness-assessing structures, these tasks open opportunities for students to share their thinking but direct them to think about specific aspects of a phenomenon instead of allowing the students themselves to determine what is relevant. In our sample, these constraints also steered the chemical thinking emphasis away from the main question, e.g., from chemical mechanism to chemical identity. To shift a thinking-constraining task to engage students more in chemical thinking, it is necessary to unpack what thinking is being constrained. In the case of teacher A's task, it could better reveal students' chemical thinking by opening ways for students to share their thinking about chemical mechanism, for example, by asking students to show representations of how they envision the arrangements of molecules changing as the puddle evaporates, and then by asking students to explain their own drawings.

Chemical-Thinking-Oriented. These task designs all required students to apply their skills and conceptual knowledge as "tools" to explain a real-world phenomenon as chemists would. Five of the FA tasks in our sample were found to have a chemical-thinking-oriented design. Teacher E, a high school chemistry teacher, designed a task to explore students' thinking about chemical causality and chemical mechanism, specifically focusing on the question, "What determines the outcomes of chemical change?" The task begins with a scenario in which three teachers, of varying sizes, each drink the same large amount of a caffeinated beverage, resulting in the death of the smallest person (Ms. X). The students are asked a series of questions about what could have caused one person to die but not the others: "How is it possible that everyone drank the same amount but only Ms. X died?" and "What do you think happened to Ms. X's body that caused her to die?" These questions are open-ended, making space for students to draw upon a wide variety of prior knowledge that they consider relevant from chemistry, biology, other sciences, or everyday experiences to answer the questions. Specifically, from chemistry, the students could apply what they had learned about ratios (amount of toxin per kilogram of body weight) and ways that some toxins can react in the body. Such tasks are likely to provide student responses that the teacher can interpret to gain insight into the students' underlying chemical thinking about chemical causality or mechanism. FAs in our sample like this had tasks and questions that aligned with one of the chemical thinking strands (questions that chemistry is uniquely positioned to answer). The tasks presented students with a scenario within reach of familiarity or which students may have experienced in their lives outside school, and then asked students to apply their chemistry knowledge to explain a phenomenon. The questions tended to be open-ended, allowing students to draw upon a wide array of prior knowledge.

Summary. Four types of designs were found in our sample. These designs ranged in their accessibility to students with tool-oriented designs being the least accessible, requiring specific prior content knowledge, and chemical-thinking-oriented designs being the most accessible, allowing students to draw on knowledge from many areas including their everyday experiences. These designs also ranged in their ability to reveal students' thinking. Some tasks presented questions with one correct answer (tool-oriented and correctness-assessing designs) while other allowed students to share their

https://dx.doi.org/10.1021/acs.jchemed.0c00361
J. Chem. Educ. 2020, 97, 4255−4267

thinking about a phenomenon (chemical-thinking-oriented design).

Purposes in Using a Formative Assessment

The analysis of the narratives about the teachers' purposes for using their FAs yielded three categories.

Conceptual Mastery. Five teachers used their FAs to assess students' understanding of specific concepts. Teacher T, when writing about the goal of using his FA, stated

> *The central goal of this assessment is the concept of concentration. The concept of concentration is very important in chemistry and all its applications.*

This teacher demonstrated awareness that the concept of concentration can be used as a tool in the application of chemistry, but his goal was to assess students' understanding of the tool rather than how it can be applied. Teacher A had a similar goal. To explain the reasons behind her FA use, she wrote

> *I designed this assessment with the goal of being able to see how much students understood a couple of things based on the misconceptions they had at the beginning of the unit.*

All of the teachers with FA purposes in the conceptual mastery category, when sharing why they chose to use the FA and the central idea that that it targeted, wrote that they identified concepts or misconceptions as the focus. These teachers were clear that their purposes for use of their FAs were to assess students' understanding of a concept and not how the students used the concepts to answer chemistry-related questions.

As with tool-oriented task designs, there are times when conceptual mastery purposes are useful. For example, teacher A's FA purpose referred to an earlier FA in her unit, so it may be that the representations she presented (see Figure 2) were based on ones that students had originally drawn. In this case, this FA may have been at the service of a larger goal that had a chemical thinking focus. This might explain the mismatch between the chemical thinking question most directly related to her FA task (chemical mechanism) and the question that was its focus (chemical identity).

Conceptual Mastery and Chemical Thinking. Four teachers communicated a purpose of assessing students' understanding of concepts along with a secondary goal of understanding how students apply concepts to explain phenomena. Their FA purposes differed from the conceptual mastery category, as these teachers included a deliberate chemical thinking focus. For example, a middle school science teacher, teacher G, wrote that her goal in using her FA at the end of a unit was to "highlight a few key concepts" about the sources of phosphorus, their relative amounts from each source, and some relevant vocabulary. Thus, part of her intent was to assess how well students had learned certain facts and vocabulary. However, she also expressed a secondary goal to

> *[see] what students are thinking about solutions for this problem [of phosphorus runoff].*

In addition to assessing the students' conceptual understanding, she wanted to learn how the students applied their understanding to control the phosphorus runoff to prevent algal blooms. This addresses the chemical control question by asking the following: How can the effects of the phosphorus runoff be controlled?

Uncovering Students' Chemical Thinking. Four teachers employed FAs with a purpose of understanding how students use their knowledge to answer questions that can be uniquely answered by chemistry. Unlike the previous category, this was not a secondary focus for these teachers; rather, it was their only stated purpose. Teacher B (a high school chemistry teacher) expressed the following about her beginning-of-unit FA:

> *I was curious to know their thoughts about what makes the three states of matter different from one another.*

This teacher wanted to investigate how students could tell the states of matter apart, rather than whether the students were bringing correct conceptions into their work. This purpose refers to the structure−property relationships question, specifically by asking the following: What cues are used to differentiate solids, liquids, and gases?

As another example, teacher E, a high school chemistry teacher, wrote

> *Students have learned enough about toxicity to make connections with chemical reactions, but the prompt is open to students' prior knowledge about the effects of caffeine, overdoses, and how the body works.*

Teacher E was interested in seeing how her students applied their chemistry knowledge to explain an overdose of caffeine. As noted earlier, her focus here is on the chemical thinking questions of chemical mechanism and chemical causality, specifically, students' thinking about what factors and processes influenced the overdose outcome. She noted that there is chemistry-specific knowledge she may expect the students to use, but she also relayed her openness to seeing what other knowledge students would draw upon. Rather than looking for students to demonstrate understanding of a concept, these teachers intended to uncover the students' reasoning about a chemistry-related phenomenon.

Summary. The teachers in our sample used FAs to assess students' conceptual mastery, to assess a mixture of conceptual mastery and chemical thinking, or specifically to uncover students' chemical thinking. Some of the teachers upheld a topic-focused emphasis, using their FAs to evaluate students' knowledge, and judging their understanding of specific content. Others blended two goals; they were concerned with the students' level of conceptual understanding but also wanted to see what resources students drew upon to explain phenomena. Others solely used a chemical thinking perspective when setting purposes for using their FA. These teachers expressed a wish to uncover what resources students drew upon to explain a phenomenon.

Evaluation Focus

The narratives created from the analysis of the teachers' notice, interpretations, and feedback on student work were analyzed to infer what appeared to be teachers' goals when evaluating the student work, rather than the goals they expressed as their intentions for the FA activity.

Conceptual Understanding. Despite the variety of task designs and expressed purposes for the FA, the vast majority (n = 10) of the teachers in our sample evaluated their students' work with a goal of assessing students' conceptual understanding. For example, when asked what would happen to phosphorus on the ground at the top of a mountain or below in the city (points A and B in Figure 3), one of teacher G's students wrote the following:

> *[On top of the mountain the phosphorus] will get into the ground in infiltration. [In the city] because there is too many houses and probably cars too... it will move to the river.*

https://dx.doi.org/10.1021/acs.jchemed.0c00361
J. Chem. Educ. 2020, 97, 4255−4267

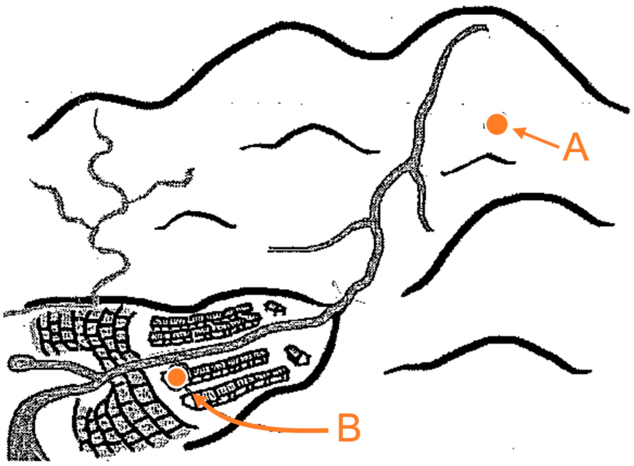


Figure 3. Representations from teacher G's FA identifying two points of interest for the students.

In response to this, teacher G wrote the following about what she noticed:

[I see] understanding of runoff and infiltration on impervious and pervious land cover, respectively.

This teacher's noticing was inferential. She did not simply rephrase that the student considered that phosphorus would runoff into the river because there are houses and cars. She inferred which concepts the student understood in order to answer in this way. In a similar way, most teachers went beyond the surface of the students' answers to infer which concepts students used correctly or misunderstood.

In response to the prompt "Draw 3 pictures of how you think the particles look in a solid, liquid, and gas below", one of teacher B's students drew clusters of three circles to represent the solid phase, clusters of two circles to represent the liquid phase, and individual circles to represent the gas phase. When evaluating the student's response, teacher B wrote

It seems the student does not realize that the water, ice, and gas have the same chemical formula and therefore the particles she drew are not the same.

The teacher interprets by judging the correctness of the student's answer and stating which knowledge the student lacks. In contrast, a sense-making interpretation based on structure−property relationships thinking behind the student's answer could be that this student may have been thinking that there are changes in the arrangements of particles within water molecules (3 particles together in solid, 2 together in liquid, and 1 alone in gas), which leads to changes in observable properties of the phases of water. Attempting to understand the student's thinking opens opportunities for a teacher to leverage productive resources in the student's thinking (e.g., in this case that structural changes occur) to advance the student's reasoning. In addition to teacher B's evaluation quoted above, this teacher also identified a correct feature of the student's drawing, that the distance between the particles increased when moving from solid to liquid to gas. This teacher, as well as others in this category of evaluation focus, tended to explain students' correct and incorrect ideas, but they did not hypothesize about the students' thinking that lay beneath responses. When interpreting student thinking and the underlying assumptions of students' answers, the teachers tended to be more corrective rather than focused on the students' sense-making. The most common ways of interpreting were to state whether students provided correct or incorrect responses, to explain which knowledge students lacked, or to remain silent on interpreting.

Most teachers did not offer hypothetical feedback as a part of their evaluations. Those who did focused on guiding students to the correct answer. A typical example of this was the following:

Can you find in your notes what the symbol M stands for? Will your answer to question 1 change now?

The goal conveyed here is for students to master skills and conceptual understanding. The teacher is not trying to advance the students' reasoning but rather concentrating on ensuring that students know the correct vocabulary, facts, and algorithms, as well as where to find these when needed.

When noticing and interpreting their students' work, these teachers made statements about which answers were correct or incorrect, and what concepts the students did and did not understand. They did not speculate about the students' underlying thinking that led to responses. Instead, they tended to focus on evaluating students' conceptual understanding. While there are certainly times that focusing on correctness is important, what stood out in the sample was the incongruence between purposes expressed for the FA (high variety) and evaluations rendered (mostly seeking to assess correctness of conceptual understanding).

Conceptual Understanding and Students' Thinking. Two teachers focused on evaluating student work for conceptual understanding. Like the teachers in the previous category, these teachers inferred which concepts the students understood, and were at times corrective, noting knowledge that the students lacked. However, these teachers differed in the type of hypothetical feedback that they posed. As an example, teacher S was interested in understanding how students think about controlling the concentration of a colored powdered drink solution. She provided the following feedback to a student:

Why did the molecules stay the same? What happened specifically to the taste? Why did the taste change?

Teacher E used a similar approach when trying to understand a student's thinking about chemical mechanism:

I would ask her what she means by that—I am wondering what that maintenance means, and if she is thinking about the way that caffeine reacts in the body. I would ask her to explain what she thinks the white blood cells are doing, and what would happen with the blood cells if there was less caffeine in [Ms. X's] body.

These questions seem to be aimed at getting the student to clarify thinking about what affects the change in concentration or how caffeine interacts in the body.

These teachers' primary strategy in giving feedback was to pose questions to prompt students to make their thinking more explicit. It is also notable that, for these teachers, eliciting students' thinking seemed to be easier than interpreting thinking in the student responses. Speculating on student thinking requires practice to overcome apprehension about the uncertainty involved in inferring.[39] Even though they did not venture to make interpretations, these teachers appeared to prioritize figuring out the thinking underneath students' conceptual understanding, which differentiated them from the previous category. Both of these teachers, when explaining how they designed their FAs, explicitly emphasized that they considered it important to offer opportunities for students to connect the task to students' own experiences beyond school.

https://dx.doi.org/10.1021/acs.jchemed.0c00361
J. Chem. Educ. 2020, 97, 4255−4267

This may be a sign that this emphasis is one that helps teachers learn about and value the usefulness of students' own thinking.

Students' Chemical Thinking. Only one teacher focused on interpreting the chemical thinking underneath students' answers. When looking at responses that students provided both before and after watching a demonstration that involved acids, teacher R remarked as follows about one student's thinking:

> *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. I wonder about whether being confronted with this particular activity helped her adjust her thinking to reflect that adding an acid might just cause a reaction, rather than "destroying" something. The idea that the acid was "breaking down" the pigment molecules was very common. I presume that students still believed by the end of the lesson that breaking things down is the primary function of an acid, and that that must be what's going on when the pigment color change takes place.*

Here, the teacher noted changes that seemed to occur in the student's thinking based on the changes in the student's language. The teacher did not focus on judging the correctness of the answers, even when they were incorrect, but instead looked to understand how the student's thinking changed, and how this could provide insights into how the student was thinking about what identifies a substance as an acid. The teacher also linked this to ways that other students were thinking, which has been found by others to be concurrent with a more inferential approach to evaluation.[18,39] Because the teacher interprets the sense-making in students' responses, she can then use ways that students think about how to identify acids to support students' development of chemical thinking (in this case, chemical identity thinking).

Summary. As with the purposes, when evaluating student work, the teachers focused on evaluating conceptual understanding or uncovering chemical thinking, or a mixture of both. Most teachers focused only on determining their students' level of conceptual understanding based on responses to the FA task. One teacher focused instead on hypothesizing what the students may have been thinking that led them to their answers. Two teachers' evaluations displayed characteristics of both of these. This latter group evaluated students' level of understanding but also sought to elicit further their students' thinking, which they demonstrated through hypothetical feedback questions.

Consistency of Focus

The goal of the year-long PD program was for the teachers to develop an FA that was focused on chemical thinking. The purpose behind our development of this analytical approach was not to see whether a teacher uses a chemical thinking approach, but to measure how and where a chemical thinking perspective grows in or fades. Most teachers' focuses changed across the FA process, from task design to intended implementation purpose for the FA to evaluation of student work (Figure 4). The Sankey diagram shows that, for many, there was weak correspondence between their stated goals and enacted evaluations. Most participants' evaluations either remained at the same extent of chemical thinking focus or decreased in chemical thinking emphasis when compared to that of their FA task design. No teachers moved more toward an increased chemical thinking focus when going from design to implementation to evaluation of student work. The only teachers with strong correspondence among their focuses, from design through evaluation, lie on the polar ends: conceptual mastery and chemical thinking emphasis. Four teachers who began with tool-oriented designs were consistent in their focus to uncover students' conceptual understanding. One teacher, on the other extreme, was also consistent, remaining focused on students' chemical thinking.

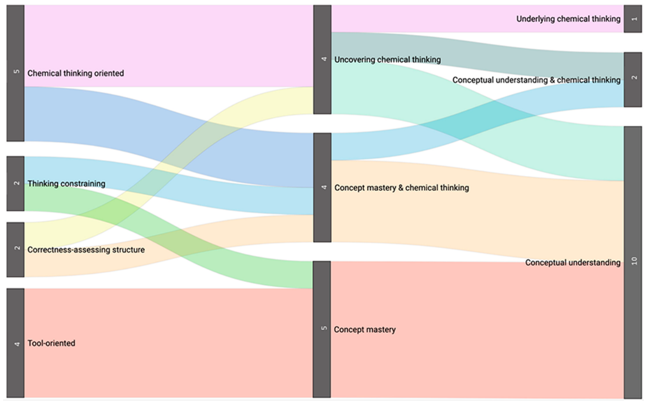


Figure 4. Numbers of teachers' FA portfolio chapters at different levels of chemical thinking across phases of FA design, teacher's purpose for the FA, and the teacher's evaluation of student work. (Note: the colors signify the number of different "paths" taken by the teacher from each node. A node with three colors means that there are three different "paths" taken by the teachers at that node.)

The weakest correspondence tended to be the shift from FA purpose to student work evaluation. There were many teachers who expressed an intent to understand students' thinking but ultimately evaluated students' work for conceptual understanding. Teachers U, G, B, and E all used chemical-thinking-oriented FA designs and had, at least, a partial goal of understanding students' thinking. However, none of these teachers ultimately interpreted students' chemical thinking in their evaluations. This may mean these teachers need greater support to develop this ability, or that they view looking for students' correct ideas as the same thing as interpreting students' thinking.[23] Teachers D, M, S, and A provided student work evaluations that may have been hindered by the same limitation, but the available student thinking was also limited by these teachers' FA task designs, similar to what was found by Harshman and Yezierski.[9]

■ CONCLUSIONS AND IMPLICATIONS

This study developed a discipline-specific approach to analyzing teachers' FA practices through the use of an FA portfolio chapter. While other studies have characterized teachers from novice to advanced, particularly on the basis of their evaluation of student work,[18] this analysis was not designed to place judgment on teachers' FA practices, but rather to characterize them in order to look for changes. The analysis was able to uncover how teachers' focuses changed throughout the FA process. Four types of FA designs were found in our sample. By the end of the year-long PD, many teachers (n = 5) had designed an FA that could reveal students' chemical thinking. Some teachers were able to develop a task that had the potential to reveal student thinking but used questions that constrained student thinking (n = 2) or

https://dx.doi.org/10.1021/acs.jchemed.0c00361
J. Chem. Educ. 2020, 97, 4255−4267

structured the FA in a way that did not allow the students to share their thinking ($n = 2$). Four of the teachers designed FAs that limited students to demonstrating content knowledge. The teachers exhibited three main purposes for their FAs, ranging from identifying which content and skills their students understand to uncovering how their students think. Similar focuses were found for the teachers when they were evaluating student work, with many more teachers ($n = 10$) adopting a conceptual understanding focus. Few teachers maintained a consistent focus, particularly from purpose to student work evaluation. Teachers whose focus changed tended to move away from a chemical thinking orientation when it came to the evaluation of student work.[17] This points to a need for bolstering teachers' ability to interpret students' chemical thinking through more practice and feedback as part of the PD.

The four teachers who used tool-oriented designs and were consistent in their focus of looking for mastery (lower red path in Figure 4) seemed to display a view of chemistry as a collection of facts that need to be learned. Evidence for this includes their sole focus being to collect and analyze student data with the aim of ensuring correct conceptual understanding. Five teachers (middle multicolor paths in Figure 4 that conclude at conceptual understanding) seemed to be in transition with regard to how they view chemistry. As we use this approach to analyze these teachers' portfolio chapters across a full year of PD, we may be able to understand better how changes over time developed. Nevertheless, these teachers designed or selected FAs that had potential to reveal students' chemical thinking and demonstrated some intent to interpret students' underlying thinking, but they focused on evaluating only conceptual understanding when examining student work. This change in focus when evaluating student work may occur for several reasons. Perhaps a teacher is not fully committed to the idea that chemistry is about the application of knowledge to analyze, synthesize, and transform matter, or a teacher may consider content knowledge and chemical thinking to be equivalent to this.[23] We have also observed in the past that teachers are uncomfortable with the uncertainty involved in hypothesizing about the assumptions that underlie chemical thinking.[40] Any of these could be evidence that teachers were in the process of expanding their FA practices to include a chemical thinking emphasis. That one teacher did maintain a chemical thinking emphasis through all three parts of the FA process also suggests that it is possible for teachers to incorporate this focus within a year of PD.

Shortening the lengthy portfolios reported in the literature[12,13,15,16] to single portfolio chapters made the data collection reasonable for teachers to complete. Teachers consistently reported that each portfolio chapter took 3−4 h to prepare. The majority also expressed appreciation that they could easily adapt their portfolio chapters to provide data for other purposes, such as teacher evaluations and evidence for licensure renewal. As noted by Harshman and Yezierski,[9] portfolio chapters as a data collection method offer rich insight into teachers' FA practices at a moment in time, and they can reveal how teachers are focusing throughout the FA process. As other researchers have done, we are also analyzing classroom videos across the year to better understand FA enactment, conducting focus group interviews to more deeply understand how teachers evaluate student work, and conducting stimulated video recall retrospective interviews with teachers to more deeply examine growth across a year, in order to develop a more robust picture of each teacher's progress during the year (findings from these have begun to be reported[39,41] and analysis is ongoing). Regarding the FA portfolio, we note that segmenting the portfolio into smaller lesson-sized pieces has limitations. This data collection approach only provides a snapshot of a teacher's FA practice and does not include enough data to illustrate the teacher's diversity of FA practices that is reported in other portfolio analyses.[12,13] However, the FA portfolio chapters do make it feasible for a teacher to provide data at multiple points throughout a year or even across a longer period of time.

FA portfolio chapters can also be useful to teachers who wish to conduct action research to better understand and strengthen their FA practice. The Supporting Information contains a guide with all examples from every category for design, purpose, and evaluation. This guide can be used by teachers to analyze their own lessons to determine how their lessons meet their own expectations, and for teachers who want to diversify their FA practices by enacting FA with a chemical thinking perspective. Teachers can also compare their FAs to those presented in this paper to examine the types of FAs they tend to use or design. Teachers may find it helpful to try designing chemical thinking FAs or to use FAs that have been designed for this purpose. Many are available at the ChemEd Xchange,[41] along with alignment to NGSS, reflections by teachers on using the FAs, and samples of student work. It is our hope the FA portfolio analysis tool offered here will support teachers who are interested in diversifying their practice to include greater chemical thinking focus.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available at https://pubs.acs.org/doi/10.1021/acs.jchemed.0c00361.

- Blank FA portfolio chapter provided to the participants (PDF, DOCX)
- Example FA portfolio chapter provided to the participants (PDF)
- Formative assessment analysis guide for teachers or researchers (PDF)
- Confirmability analysis (PDF)

■ AUTHOR INFORMATION

Corresponding Author

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

Author

Timothy N. Abell − *Department of Chemistry, University of Massachusetts Boston, Boston, Massachusetts 02125, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jchemed.0c00361

Notes

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This material is based on work supported by the National Science Foundation award DRL-1621228.

■ REFERENCES

(1) Bell, B.; Cowie, B. The Characteristics of Formative Assessment in Science Education. *Sci. Educ.* **2001**, *85* (5), 536−553.
(2) 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), 799−824.
(3) Black, P.; Wiliam, D. Assessment and Classroom Learning. *Assess. Educ. Princ. Policy Pract.* **1998**, *5* (1), 7−74.
(4) 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.
(5) Richards-Babb, M.; Curtis, R.; Georgieva, Z.; Penn, J. H. Student Perceptions of Online Homework Use for Formative Assessment of Learning in Organic Chemistry. *J. Chem. Educ.* **2015**, *92* (11), 1813−1819.
(6) Emenike, M.; Raker, J. R.; Holme, T. Validating Chemistry Faculty Members' Self-Reported Familiarity with Assessment Terminology. *J. Chem. Educ.* **2013**, *90* (9), 1130−1136.
(7) Raker, J. R.; Holme, T. A. Investigating Faculty Familiarity with Assessment Terminology by Applying Cluster Analysis to Interpret Survey Data. *J. Chem. Educ.* **2014**, *91* (8), 1145−1151.
(8) Underwood, S. M.; Posey, L. A.; Herrington, D. G.; Carmel, J. H.; Cooper, M. M. Adapting Assessment Tasks to Support Three-Dimensional Learning. *J. Chem. Educ.* **2018**, *95* (2), 207−217.
(9) 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.
(10) 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.
(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) 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.
(13) 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.
(14) 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.
(15) Baxter, G. P.; Bass, K. M.; Glaser, R. *An Analysis of Notebook Writing in Elementary Science Classrooms (CSE Report 533)*; University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing (CRESST): Los Angeles, 2000. http://cresst.org/publications/cresst-publication-2905/ (accessed 2020-10-01).
(16) 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.
(17) 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.
(18) 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.
(19) 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.
(20) 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.
(21) Hattie, J.; Timperley, H. The Power of Feedback. *Rev. Educ. Res.* **2007**, *77* (1), 81−112.
(22) 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.
(23) Morrison, J. A.; Lederman, N. G. Science Teachers' Diagnosis and Understanding of Students' Preconceptions. *Sci. Educ.* **2003**, *87* (6), 849−867.
(24) Sevian, H.; Talanquer, V. Rethinking Chemistry: A Learning Progression on Chemical Thinking. *Chem. Educ. Res. Pract.* **2014**, *15* (1), 10−23.
(25) National Research Council. *A Framework for K-12 Science Education*; National Academies Press: Washington, DC, 2012.
(26) Ngai, C.; Sevian, H. Capturing Chemical Identity Thinking. *J. Chem. Educ.* **2017**, *94* (2), 137−148.
(27) 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.
(28) Stains, M.; Escriu-Sune, M.; Molina Alvarez de Santizo, M. L.; 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.
(29) Talanquer, V. Progressions in Reasoning about Structure-Property Relationships. *Chem. Educ. Res. Pract.* **2018**, *19* (4), 998−1009.
(30) 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.
(31) 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.
(32) 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*, 3478−3489.
(33) 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.
(34) 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.
(35) Borko, H.; Stecher, B. M.; Alonzo, A. C.; Moncure, S.; Mcclam, S. Artifact Packages for Characterizing Classroom Pratice: A Pilot Study. *Educ. Assess.* **2005**, *10* (2), 73−104.
(36) Ambitious Science Teaching. Eliciting Students' Ideas. https://ambitiousscienceteaching.org/eliciting-students-ideas-2/ (accessed 2020-10-01).
(37) Ruiz-Primo, M. A.; Li, M. Analyzing Teachers' Feedback Practices in Response to Students' Work in Science Classrooms. *Appl. Meas. Educ.* **2013**, *26* (3), 163−175.
(38) 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.
(39) Szteinberg, G.; Balicki, S.; Banks, G.; Clinchot, M.; Cullipher, S.; Huie, R.; Lambertz, J.; Lewis, R.; Ngai, C.; Weinrich, M.; et al. Collaborative Professional Development in Chemistry Education

Research: Bridging the Gap between Research and Practice. *J. Chem. Educ.* **2014**, *91* (9), 1401−1408.
(40) Dini, V.; Sevian, H.; Caushi, K.; Orduña Picón, R. Characterizing the Formative Assessment Enactment of Experienced Science Teachers. *Sci. Educ.* **2020**, *104* (2), 290−325.
(41) Assesing for Change in Chemical Thinking. www.chemedx.org/ACCT (accessed 2020-10-01).

https://dx.doi.org/10.1021/acs.jchemed.0c00361
J. Chem. Educ. 2020, 97, 4255−4267