



Chemistry  
Education Research  
and Practice

**Investigating High School Chemistry Teachers' Assessment  
Item Generation Processes for a Solubility Lab**

Journal:	<i>Chemistry Education Research and Practice</i>
Manuscript ID	RP-ART-04-2020-000121.R2
Article Type:	Paper
Date Submitted by the Author:	20-Oct-2020
Complete List of Authors:	Schafer, Adam; Miami University, Chemistry & Biochemistry Yeziarski, Ellen; Miami University, Chemistry & Biochemistry

SCHOLARONE™  
Manuscripts

## ARTICLE

# Investigating High School Chemistry Teachers' Assessment Item Generation Processes for a Solubility Lab

Adam G. L. Schafer<sup>a</sup> and Ellen J. Yeziarski<sup>a</sup>

Received 00th January 20xx,

Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Designing high school chemistry assessments is a complex and difficult task. Although prior studies about assessment have offered teachers guidelines and standards as support to generate quality assessment items, little is known about how teachers engage these supports or enact their own beliefs into practice while developing assessments. Presented in this paper are the results from analyzing discourse among five high school chemistry teachers during an assessment item generation activity, including assessment items produced throughout the activity. Results include a detailed description of the role of knowledge bases embedded within high school chemistry teachers' pedagogical content knowledge and the processes used to enact these knowledge bases during planned formative assessment design. Implications for chemistry teacher professional development are posited in light of the findings as well as potential future investigations of high school chemistry teacher generation of assessment items.

## 1 Introduction

Many states and school districts are implementing reform educational practices aimed at helping students build conceptual understanding of chemistry content and gain skills thinking like a scientist. Assessment has become an increasingly critical part of these reforms because the practices guiding assessment will ultimately be used to evaluate the success of the reform practices (Edwards, 2013; National Research Council, 2014). However, little research has been conducted on the process teachers undergo to design assessments, especially chemistry specific assessment considerations at the high school level. Current literature provides teachers with guidelines for designing assessments and how these guidelines impact the interpretability of assessment items by students and assessment results by teachers (Bell and Cowie, 2001; Towndrow *et al.*, 2010; Ruiz-Primo *et al.*, 2012; Towns, 2014; Harshman and Yeziarski, 2014; Dini *et al.*, 2020). If teachers could effectively enact these guidelines into practice, teachers nationwide would demonstrate high-quality assessment practices; however, several relevant studies reveal this is not the case (Park and Oliver, 2008; Sandhu *et al.*, 2015; Cisterna and Gotwals, 2018; Siegel *et al.*, 2019). A thorough understanding of current assessment practices used by chemistry teachers to develop assessments is necessary to better help teachers develop their ability to assess student knowledge and gauge the success of learning activities.

## Formative Assessment Design Practices

Formative assessments are used by teachers to inform continued learning (Black and Wiliam, 1998; Irons, 2008).

Clinchot *et al.*, 2017). Responses to in-class questioning and planned quizzes are a few examples of the results of formative assessments that provide teachers and students with much needed information about the progress of learning and the success of the learning activities. Designing planned formative assessments is a complex process that involves a multitude of decisions that ultimately impact the teacher's ability to draw conclusions about the teaching and learning process (Bell and Cowie, 2001; Remesal, 2011). The opportunity to make informed decisions about the students, learning activities, or assessment quality is dependent upon decisions that the teacher makes while planning and generating formative assessments. A teacher's understanding of formative assessment and the practices they employ depends on their ideas about intelligence, the process of teaching and learning, the nature of assessment tasks, and evaluation criteria (Wolf *et al.*, 1991). In addition, the wide variety of teacher approaches and understanding of assessment causes students to experience classroom cultures that value and use assessments in significantly different ways (Shepard, 2000). In the end, teachers must not only apply their conceptual understanding of the topic(s) of interest but also a practical knowledge of their students and the classroom environment when designing planned formative assessments if they hope to inform day-to-day instruction (DeLuca *et al.*, 2018).

Formative assessment practices have been the focus of several domain-general investigations (Tomanek *et al.*, 2008). Many formative assessment practices can be applied without domain-specificity, for example, all teachers can provide their students with opportunities to demonstrate knowledge and skills that can be used to inform continued learning. However, specific formative assessment practices employed by teachers are highly dependent upon expertise and cultures embedded within the discipline (Coffey *et al.*, 2011). For example, chemistry teachers need a wealth of knowledge about chemical species and how

<sup>a</sup> Miami University, Department of Chemistry and Biochemistry, Oxford, OH, USA.

E-mail: yeziere@miamioh.edu

those species interact to effectively design a planned format  
assessment about the solubility of ionic salts that appropriate  
evokes student responses that will be informative for guidance  
continued instruction.

In the absence of chemistry-specific support, chemistry  
teachers are left translating the available domain-general  
guidelines into practices suitable for their own classrooms (Black  
and Wiliam, 1998). As the Next Generation Science Standards  
(NGSS) become more common and widespread, teachers  
being provided more resources for designing formative  
assessments that are 3-dimensional (encompass the NGSS  
disciplinary core ideas, cross-cutting concepts, and science  
engineering practices) (NGSS Lead States, 2013; Harris *et al.*  
2016; Underwood *et al.*, 2018). However, enacting the reform  
formative assessment practices of the NGSS effectively can be  
difficult; the NGSS could provide additional challenges for  
teachers to overcome. A recent report from the National  
Academies of Science, Engineering, and Medicine emphasizes  
the importance of identifying core practices for teachers to  
develop to meet the demands of NGSS (National Academies of  
Sciences, Engineering, 2015). The report specifically  
communicated that “teachers need to master a range of formative  
and summative assessment strategies” (p. 103) when examining  
assessment’s critical role in informing instruction.

The effective use of formative assessment is linked to  
improved student engagement and performance (Ruiz-Primo  
Furtak, 2007; Furtak *et al.*, 2016). However, teachers may find  
translating domain-general support for implementing formative  
assessments difficult and newly implemented practices may  
result in student outcomes that are not measurable (Buck  
Trauth-Nare, 2009; Gómez and Jakobsson, 2014; Harshman  
Yeziarski, 2017). Additionally, very few investigations study  
how teachers interact with specific content during assessment  
design (Tomanek *et al.*, 2008; Coffey *et al.*, 2011). Often  
research about assessment practices focuses on general  
knowledge and beliefs about assessment (Remesal, 2011; O’Neil  
2015; Yan and Cheng, 2015). common assessment tools (Black  
and Wiliam, 1998; Suskie, 2009), and grading practices  
(Henderson *et al.*, 2004; Dubey and Geanakoplos, 2010; Toland  
and Dubas, 2017) but not on teachers’ assessment design  
process, especially chemistry teacher assessment design  
processes (Coffey *et al.*, 2011).

One study by Tomanek and others investigated teacher  
considerations when selecting and evaluating formative  
assessment tasks (Tomanek *et al.*, 2008). In the study by  
Tomanek *et al.*, several prospective and practicing teachers  
selected assessment tasks or evaluated the potential of a task to  
assess student understanding (Tomanek *et al.*, 2008). Their  
findings show that teachers exhibit general tendencies when  
evaluating/selecting assessment items, often influenced by two  
broad categories of concerns: (1) characteristics of the task and  
(2) characteristics of the students or the curriculum (Tomanek *et al.*  
2008). Their findings align with other calls for further  
investigation into teachers’ process of developing assessments  
(Coffey *et al.*, 2011; Park and Chen, 2012). The process of  
designing assessment items is a personal experience (Yan and  
Cheng, 2015), and a study capturing teachers’ considerations

during this process may expand the understanding about how  
teacher beliefs about assessment are translated into practice.

## Frameworks

The assessment design process weaves together many  
fragments of a teacher’s knowledge. When designing  
assessments, a teacher must not only consider the content to  
assess, but also how to construct assessment items to best elicit  
student knowledge about the content. What a teacher considers  
when designing assessments is informed by their pedagogical  
content knowledge (PCK) (Marzabal *et al.*, 2018). An early  
definition of PCK from Shulman described it as knowledge  
“beyond subject matter per se to the dimension of subject matter  
knowledge for teaching” (Shulman, 1986). Essentially, PCK  
informs the decisions teachers make about what concepts to  
assess, what kinds of problems to assign, the instruments used to  
elicit student ideas, and how to interpret student responses  
(Marzabal *et al.*, 2018; Hume *et al.*, 2019). Many models  
describing PCK have been proposed since its conception  
(Magnusson *et al.*, 1999; Park and Oliver, 2008; Hume *et al.*,  
2019). Recently, members of the 2<sup>nd</sup> PCK summit gathered to  
generate the Refined Consensus Model of PCK (Hume *et al.*,  
2019). In this model, PCK is defined as “the knowledge of,  
reasoning behind, and planning for teaching a particular topic in  
a particular way for a particular purpose to particular students for  
enhanced student outcomes” and is characterized through the use  
of five knowledge bases that are described in Table 1 (Gess-  
Newsome, 2015; Hume *et al.*, 2019).

The knowledge bases in Table 1 describe the collective  
knowledge available to inform the processes of teaching a  
particular topic in a particular way for a particular purpose  
(Hume *et al.*, 2019). Early conceptualizations positioned PCK as  
a separate knowledge domain alongside the knowledge bases  
(Shulman, 1987). More recent views recognize that PCK is not a  
“freestanding type of knowledge,” although one’s PCK is  
continuously influenced by and influences the embedded  
knowledge bases (e.g., Magnusson *et al.*, 1999; Abell, 2008;  
Park and Chen, 2012). As such, evidence of one knowledge base  
does not “equate to PCK.” However, we posit that our  
understanding of the complex nature of PCK can be strengthened

Table 1. Pedagogical Content Knowledge Bases as defined by the Refined Consensus Model (Gess-Newsome, 2015; Hume *et al.*, 2019)

Knowledge Base (Knowledge of...)	Description (Knowledge of...)
Content	the academic ideas and concepts that are pertinent to a discipline
Curriculum	the structure, scope, sequence, and goals of a curriculum
Students	students’ general characteristics, cognitive development, and variations in approaches to learning
Pedagogy	skills and strategies related to learning theories, instructional principles, and classroom management
Assessment	the design and interpretation of formative and summative assessments as well as how to take action from assessment data

by better understanding how the integrated components (i.e., knowledge bases) are enacted to transform knowledge opportunities for learning.

Each of the knowledge bases in Table 1 can be described existing within the collective community at large, within the mind of the individual, or be enacted by the individual (Park and Oliver, 2008; Hume *et al.*, 2019). A teacher may have access to knowledge from a particular knowledge base (collective-PCK) but may not hold that knowledge personally (personal-PCK) or enact it in their classroom practices (enacted-PCK) (Park and Oliver, 2008; Hume *et al.*, 2019). The five knowledge bases do not exist in isolation and are highly interconnected (Park and Chen, 2012; Hume *et al.*, 2019). For example, a chemistry teacher would likely consider ideas within “knowledge of content” when designing and implementing an assessment (which would require the application of “knowledge of assessment”). Although enacted-PCK is typically used to describe in-class activities only, a teacher’s knowledge applied to activities outside of class is likely not drastically different. For example, a teacher will still enact some knowledge about assessments when designing formative assessments or interpreting assessment results. As such, providing teachers an environment that encourages them to communicate ideas during generation of planned formative assessment items could reveal teacher practices for translating their personal-PCK into enacted chemistry assessment knowledge.

## Research Questions

The purpose of this investigation is to identify the role of PCK and characterize the processes used by teachers to enact their PCK during the design of planned chemistry formative assessments. The research questions guiding this investigation are:

1. What is the role of high school chemistry teachers’ pedagogical content knowledge when generating planned formative assessment items for a solubility lab?
2. What processes do high school chemistry teachers undergo when enacting their pedagogical content knowledge when designing planned formative assessment items for a solubility lab?

## Methods

To address the research questions, a group of high school chemistry teachers were observed while participating in an assessment item generation activity. The activity was part of an IRB-approved long-term professional development (PD) investigating the alignment between assessment beliefs and practices.

## Sample

Five public high school chemistry teachers participated in this study. All participating teachers were previous members of the Target Inquiry at Miami University (TIMU) project and were familiar with the inquiry lab used for the assessment item

generation activity (Yeziarski and Herrington, 2011). All teachers taught in public rural and suburban schools, ranging in size from 400 – 1,600 students at the time of the PD.

## Item Generation Activity

Teachers worked in two groups, separately generating one formative assessment item for each of the educational objectives (EOs) of an inquiry lab titled *What’s the Solution?* (available on the TIMU website) (Target Inquiry at Miami University). At the time of the PD, the *What’s the Solution?* lab did not have published corresponding assessment items. Anne (11 years of experience) and Celine (18 years of experience) worked together in Group 1, while Group 2 consisted of Ashton, Claude, and Emmerson (25, 16, and 10 years of experience, respectively). The materials available for implementors on the website include the EOs, content addressed, misconceptions targeted, and prior knowledge expected of students. Both groups of teachers were audio-video recorded while generating items, and photographs of the items generated by the teachers were collected. By requiring teachers to collaborate for assessment item generation, ideas could be captured as teachers communicated to each other. Teachers completed item generation after about 20 minutes. Although not part of this investigation, teachers then critiqued the items of their peers, discussing positive item characteristics and potential improvements as a latter part of the PD. Audio files were transcribed verbatim and deductively coded using the PCK knowledge bases from the Refined Consensus Model as categories (Hume *et al.*, 2019). Afterward, statements within each category were further coded, inductively, using constant comparative analysis (Maxwell, 2013).

Excerpts of teacher dialogue were coded according to considerations expressed during item development that were classified within PCK categories. For this study, a “consideration” is an idea communicated by the teachers during assessment item generation. Enacted-PCK describes when a teacher’s personal-PCK is applied to a teaching and learning situation. This could arguably occur in the classroom environment or outside the classroom in situations such as designing an assessment or interpreting assessment results. By observing chemistry teachers during assessment item generation for a single lab, the study described herein focuses on enacted-PCK (i.e., the PCK that teachers actually employ) (Park and Oliver, 2008; Hume *et al.*, 2019). Furthermore, since the teachers were observed generating assessment items for a solubility lab, the study focuses on enacted-PCK specific to the chemistry topic of solubility and any related content (Magnusson *et al.*, 1999; Abell, 2008) as characterized through the knowledge bases in Table 1.

To evidence trustworthiness of the coding, frequent interrater debriefings were held between two members of the research team. During the debriefings, two raters separately coded roughly 15% of the teacher considerations during assessment item generation as a means of establishing interrater agreement. Additionally, code applications were presented to graduate students and chemistry education researchers who were not affiliated with this investigation. Comparison of code

Table 2. Teacher-Generated Assessment Items

EO	Group 1 Assessment Items	Group 2 Assessment Items
1	Which choice below correctly expresses what happens to the sodium particles in NaCl when sodium chloride dissolves in H <sub>2</sub> O? a. Na b. Na <sup>+</sup> c. NaH <sub>2</sub> O d. None of the above	Ca(NO <sub>3</sub> ) <sub>2</sub> (s) dissolves in water. a. What will you observe? b. Write the equation that describes this process. c. Draw a particulate model of the Ca(NO <sub>3</sub> ) <sub>2</sub> (aq) after it is all dissolved.
2	If the polarity of the particles in H <sub>2</sub> O is switched so that the “O” end is now partially positive, and the “H” end is now partially negative, what part of the NaCl would be attracted to the “H” end of the water molecule? a. “H” end of water would surround the Na ions of the salt crystal b. “H” end of water would surround the Cl <sup>-</sup> ions of the salt crystal c. “H” end to “O” end of water d. None of previous	Draw a calcium ion in water include 4 water molecules in your drawing. Explain why they are arranged in the way you provide.
3	What is the difference in charge of the Cl particle before and after NaCl dissolves in water? a. Before its negative, after its neutral b. Before its negative, after its negative c. Before its neutral, after its neutral d. Before its positive, after its negative	Na <sub>2</sub> SO <sub>4</sub> + H <sub>2</sub> O → a. Na <sub>2</sub> <sup>+2</sup> (aq) + SO <sub>4</sub> <sup>-2</sup> (aq) b. 2Na(aq) + SO <sub>4</sub> (aq) c. 2Na <sup>+</sup> (aq) + SO <sub>4</sub> <sup>-2</sup> (aq) d. 2Na(aq) + S(aq) + 2O <sub>2</sub> (g)

applications from the two coders resulted in 98% agreement. 285  
 exceptionally high interrater agreement was likely a result 286  
 frequent debriefings during the code generation process. 287  
 disagreement in code applications was discussed by the raters 288  
 followed by minor modification of code descriptions with 289  
 necessary. This iterative process of consensus building continued 290  
 with reapplication of codes to the data set and interrater 291  
 debriefing until complete agreement was reached. 292  
 The coded statements were used to characterize high school 293  
 chemistry teachers' processes for generating planned format 294  
 chemistry assessment items. By matching the audio and video 295  
 files, teacher ideas were linked to the item being generated at 296  
 time the idea was communicated. 297

Table 3. Ideas Communicated During Item Generation

Idea Communicated	Knowledge Base (Knowledge of...)
Student Prior Knowledge	Students
Student Response	Students
Representational Level	Content
Dissolving	Content
Substance	Content
Electrostatic Interaction	Content
Item Format	Assessment
Learning Activity (Lab)	Curriculum
Educational Objective	Curriculum
State/National Standards	Curriculum

## Results and discussion

### Teacher-Generated Assessment Items

The Educational Objectives (EOs) for the *What's the Solution?* 316  
 inquiry lab are: 317

1. The student will be able to correctly predict what happens when a soluble salt dissolves in water.
2. The student will be able to explain how an ionic compound dissolves in water, noting the role of water molecules in this process.
3. The student will be able to explain what happens to the charge of the ions during solvation and be able to explain why.

The EOs are stated in both the student and teacher versions, available on the TIMU website (Target Inquiry at Miami University). Each group of teachers generated one assessment item per EO, shown in Table 2.

### PCK Ideas Communicated During Item Generation

To address the first research question, ideas communicated by the chemistry teachers during assessment item generation were identified and classified by PCK knowledge base. Table 3 shows the codes representing the ideas communicated and the knowledge base that best aligns to each code. Descriptions of each of the ideas in Table 3 are provided in Appendix A.

Following Table 3 is a brief description of the ideas communicated within each knowledge base.

**Knowledge of Pedagogy.** Missing from the results in this investigation are statements within the “knowledge of pedagogy” knowledge base. However, the lack of teacher statements in this knowledge base is not surprising, since the “knowledge of pedagogy” knowledge base includes general considerations about teaching and classroom management while the teachers in this investigation were focused on generating assessment items for a specific chemistry topic.

**Knowledge of Students.** The “knowledge of students” knowledge base was the most frequent consideration of the

teachers. Ideas in this category were coded as either student knowledge or student response. When considering student knowledge, the teachers communicated the knowledge expected the students to have when responding to the assessment item. For example, teachers considered students' familiarity with constructing certain types of responses (Ashton: *My concern is that part of it will be if this is the first time they ever had to draw anything that's two separate lessons there.*) information critical for understanding the item (Celine: *If I don't remember [water's polarity] it doesn't matter because telling them.*)

Teacher considerations about expected student answers for a question were coded as student response (Anne: *A lot of them have a tendency to pick none of the above.*) and elements of student response that the teachers were attempting to elicit (Celine: *What I'm getting at is would the kid be able to tell that the oxygen end would be attracted to the chlorine now. Because the oxygen is positive now.*)

These findings align with the literature about the "knowledge of students" knowledge base. For example, Magnusson et al. state that teachers' "knowledge of students" includes considerations about the knowledge required for learning (in this case assessing) and knowledge of areas of student difficulty (Magnusson et al., 1999). The teachers' considerations about student responses included ways to construct the assessment item that either addressed common student difficulties to avoid confusion (Anne: *Would you give them the original water and polarity? Just so that's not what they miss.*) or employed them to generate purposeful item distractors (Anne: *Yeah, since students] would all choose 'none of the previous' because where's the chloride?*)

**Knowledge of Content.** The "knowledge of content" knowledge base includes all considerations about the subject matter to be assessed. Shulman recognized that to properly employ content knowledge, teachers needed to move past basic facts and concepts to understand what content is necessary warranted for a given topic (Shulman, 1986). Related literature about PCK has expanded this notion to specify the difference between topic-specific PCK and general-PCK (Magnusson et al., 1999; Park and Oliver, 2008; Hume et al., 2019). There is evidence to suggest teacher content knowledge impacted the quality of the items generated. (e.g., Celine: *The hydrogen of the water molecule would surround the sodium end of the crystal. The sodium whatever I don't know the correct terminology here.*) However, the study is bounded in characterizing the role of PCK during item generation and processes undergone by teachers to enact PCK, not on the quality of PCK or items generated. Later, the teachers who authored the item had the opportunity to engage in peer critique, which led to minor modifications to the items that are not presented here. Teachers in this investigation generated assessment items for a lab about salts dissolving in solution. As such, the nature of the PD activity limited much of the teachers' considerations to ideas about solubility. When communicating ideas about content, teachers frequently discussed not only what chemical species to include (substance, Claude: *Instead of calcium bromide, I'd use calcium nitrate*), but also the electrostatic interactions to address

in the assessment item (Ashton: *the somewhat negative portion of oxygen with the positive ion*). Throughout discussion, teachers frequently revisited the "knowledge of content" knowledge base to narrow the content addressed in the assessment items being generated (Claude: *What is it about water to dissolve in the first place?*).

In addition to the topic-specific content, teachers evoked chemistry-general ideas within the "knowledge of content" knowledge base by considering how to represent the content in the assessment items being generated. For example, teachers in this investigation were familiar with Johnstone's representational levels (Johnstone, 1991), and communicated how the representational level was important in defining how the content was addressed in the assessment item (Claude: *So, correctly predict symbolically and particulate what happens. Don't we have to include macroscopically as well?*) In this regard, the "knowledge of content" knowledge base served to inform not only what content was addressed in the assessment items, but also how that content would be perceived by the student taking the assessment.

**Knowledge of Assessment.** The "knowledge of assessment" knowledge base, although frequently communicated, did not result in a variety of codes. For the item generation PD activity, teachers were asked to generate formative assessment items. Formative assessment items are those which are used to inform continued learning after assessment (Black and William, 1998; Irons, 2008; Clinchot et al., 2017). The awareness that items generated were to be able to be used in a formative manner likely impacted teachers' assessment item design process. Throughout the generation of assessment items, teachers actively considered the method of evaluating student competency by communicating the item format (Celine: *We could make it into a multiple-choice question.*) However, the reasoning communicated by the teachers for the item formats they employed was often sourced from other knowledge bases. For example, when Group 2 was discussing what item format to use for the EO2 assessment item, they opted not to use a multiple-choice format due to considerations from the "knowledge of students" knowledge base (Claude: *I mean a multiple-choice question ... [the student] may look into it and just know factually based on seeing, but that's not necessarily explaining it.*) Claude's consideration of student knowledge with the content justified not using a multiple-choice format. In this regard, the "knowledge of assessment" knowledge base served to inform the methods of eliciting evidence of student competency, but the other knowledge bases provided insight on the affordances and hindrances of using that method. Using the "knowledge of students" to inform considerations about the "knowledge of assessment" demonstrates the interconnectedness of the PCK knowledge bases.

**Knowledge of Curriculum.** The interconnectedness of the knowledge bases was likewise instrumental in defining the role of the "knowledge of curriculum" knowledge base. When communicating ideas in the "knowledge of curriculum" knowledge base, teachers generally considered how EO and learning activity elements could inform the generation of the assessment item. Although other investigations into PCK reveal

that knowledge of specific goals and objectives as well as knowledge of general curricular structure both inform “knowledge of curriculum” knowledge base, the top-down specificity of the PD activity likely encouraged teachers to consider specific EO and lab elements essential for assessment (Magnusson *et al.*, 1999; Hume *et al.*, 2019). As such, teachers communicated ideas about the learning activity and frequently, with the rare mention of state-level objectives. Teachers attempted to ensure the assessment item similar to the lab in regard to the student response, such as having the student perform similar tasks as in the lab. Additionally, teachers sought to ensure similarity in the chemical species used in the assessment item by incorporating substances that are equally common and complex as those used in the lab, within the context of their classrooms. For example, part of the lab involved dissolving copper (II) chloride in water; however the teachers decided not to use copper (II) chloride in their assessment items (Claude: *I wouldn't use chloride because chloride is what's in [the lab].*) Instead, the teachers opted to find salts that were similar to incorporate in their assessment items (Celine: *[The lab] never uses sodium chloride. So, let's use sodium chloride as an assessment question.*) Essentially, the teachers in this investigation employed their “knowledge of curriculum” to connect their “knowledge of content” to their “knowledge of assessment.”

Similar to their considerations of the learning activity, teacher ideas communicated about the EOs bridged the gap between their “knowledge of content” and “knowledge of assessment.” Teachers consistently referred to the EOs to establish what content to address in the assessment item (Celine: *What [the students] are doing is applying the opposites attract, and that water is pulling at parts [of the salt] which is the objective.*), how that content should be represented (Claude: *The expression ‘what happens’ is essentially covering all three [representational] levels.*), and what to ask the student to do when responding to the item (Celine: *...can they still apply the idea?*)

Analyzing teacher discourse while identifying quotations that reveal teachers’ “knowledge of curriculum” demonstrated the interconnectedness of the PCK knowledge bases. Literature sources argue that the knowledge bases serve as a heuristic device, allowing for the representation of knowledge in the mind

of the teacher and that teachers likely activate multiple knowledge bases simultaneously (Shulman, 1986; Tamir, 1991; Magnusson *et al.*, 1999; Park and Oliver, 2008; Abell and Siegel, 2011; Hume *et al.*, 2019). In this investigation, the teachers’ wove together PCK knowledge bases during assessment item generation to inform processes of checks and balances necessary for connecting the instructional materials (i.e., lab and EOs) with the assessment items. These “processes” represent the transformation of personal PCK into learning opportunities as characterized by the knowledge bases and are explored further in the next section through the lens of discussion vignettes.

### Teacher Processes for Enacting PCK

In this investigation, chemistry teachers’ PCK knowledge bases served the role to inform processes that occur during the generation of planned formative assessment items. These processes involved establishing:

1. an appropriate task for the student to accomplish.
2. the content to assess versus what content is provided to the student.
3. an appropriate representation of the content in the stem and in the student response.

The vignettes in Tables 4, 5, 6, and 7 show brief exchanges between the teachers during the assessment item generation activity. For reference, each vignette includes the assessment item generated from the discussion. The vignette in Table 4 represents the teachers’ process for establishing an appropriate task during the item generation process. The “task” of an assessment item is defined as the work to be completed by the learner to demonstrate competency of what was to be learned (McDonald, 1964; Hoffman and Medsker, 1983; Jonassen *et al.*, 1999; Merrill, 2007).

In the Table 4 vignette Ashton initially proposed having students perform a multiple-choice task to assess their knowledge of how an ionic compound dissolves in water. Afterwards, the teachers collectively reasoned with the proposed task by considering if a multiple-choice task effectively assesses the students’ ability to “explain,” as is stated in the EO. Throughout this exchange, the teachers communicated ideas from several PCK knowledge bases. For example, the “knowledge of content” knowledge base was communicated as teachers reasoned with the chemical

Table 4: Ashton, Emerson, Claude item generation EO2

Statements	Assessment item
Ashton: <i>Now for this one I have an idea. What if we have three drawings showing water around an ion; one of them is correct.</i>	Draw a calcium ion in water include 4 water molecules in your drawing.
Claude: <i>I love it.</i>	Explain why they are arranged in the way you provide.
Ashton: <i>You know the somewhat negative portion of the oxygen with the positive ion, and then one of them where they are just totally wrong. So, how about this one drawing something. I don't know I'm just thinking.</i>	
Emmerson: <i>So, does that explain though?</i>	
Claude: <i>It's a good question. And specifically, what is it about water to dissolve in the first place? Right. I mean a multiple-choice question where you have the three different possible orientations you may look into it and you may just know factually based on seeing but that's not necessarily explaining it... How about how about this. Instead of a multiple-choice, draw calcium ion and four water molecules and explain why you arranged them the way that you did.</i>	

Table 5: Anne and Celine item generation EO2

Statements	Assessment item
Celine: <i>So, if the water's polarity was reversed, the hydrogen end was negative and the oxygen end was positive, which choice would describe water's interaction with salt? So, let me tell you what I'm trying to get at.</i>	<p>If the polarity of the particles in H<sub>2</sub>O is switched so that the "O" end is now partially positive, and the "H" end is now partially negative, what part of the NaCl would be attracted to the "H" end of the water molecule?</p> <p>a. "H" end of water would surround the Na ions of the salt crystal</p> <p>b. "H" end of water would surround the Cl<sup>-</sup> ions of the salt crystal</p> <p>c. "H" end to "O" end of water</p> <p>d. None of previous</p>
Anne: <i>Are you getting at solvation?</i>	
Celine: <i>Yeah. Because it's like we give the kid a situation in which we say water is this now. Okay?</i>	
Anne: <i>Okay</i>	
Celine: <i>If water's polarity was this way first. Explain which choice would correctly explain how water would interact with sodium chloride. What I'm getting at is would the kid be able to tell me that the oxygen end would be attracted to the chlorine now because the oxygen is positive now.</i>	
Anne: <i>So, you're getting at whether they know that chlorine is going to be negative.</i>	
Celine: <i>Right. So, the idea is noting the role of water molecules in this process. They know that this is switched, but can they still apply the idea that the positive end of water.</i>	
Anne: <i>So, would you give them the original water and its polarity? Just so that's not what they miss. Is not remembering [the polarity of water].</i>	
Celine: <i>No. I wouldn't. If they don't remember [water's polarity] it doesn't matter because I'm telling them imagine if water's polarity was reversed and the oxygen is now positive and the hydrogen end is now negative.</i>	
phenomena in the item (Ashton: You know the somewhat negative portion of the oxygen with the positive ion...)	
"knowledge of assessment" when considering the elicitation students' ability to "explain" (Emmerson: So, does that explain though?). While developing this item, the teachers kept originally proposed chemical phenomenon, but wove together ideas from multiple knowledge bases to generate a task for student that (they believed) aligned to the requirements stated the EO. These findings agree with other investigations that similarly identified the importance teachers place on establishing an appropriate task for the student to perform when evaluating competency (Tomanek et al., 2008). Essentially, a chemistry teacher must establish if "what the student does" appropriately evaluates student competency within a particular chemistry topic when judging competency. The process of establishing a task represents the translation of personal-PCK into enacted-PCK assessment item generation (Hume et al., 2019).	Tables 6 and 7 illustrate teachers' process for establishing how information should be represented in the assessment item. Establishing how information should be represented was an important process for teachers while generating assessment items. In the Table 6 vignette, the teachers grappled with language in the EO that set the requirements for how to represent the data (Claude: <i>The expression 'what happens' that's essentially covering all three [representational] levels.</i> ) The teachers considered ideas from the "knowledge of content" (Ashton: <i>What if we just say calcium chloride plus water with the model and a symbol?</i> ) and "knowledge of assessment" (Claude: <i>It could be three parts, right?</i> ) knowledge bases while establishing how information should be represented. When considering how information should be represented, the teachers often referred to Johnstone's representational levels as a framework to guide their design (Johnstone, 1991).
The vignette in Table 5 illustrates the teachers' process establishing what content to assess versus what content to provide to the student.	
In the Table 5 vignette Celine proposes a hypothetical "switch" of water's polarity in order to provide the student a novel situation using a familiar chemical phenomenon. Following Celine's initial proposal, the teachers collaboratively considered if the content is appropriately matched to the EO as well as what supporting content is necessary for the student to be able to answer the question. Establishing the content to assess another translation of the teacher's personal-PCK into enacted-PCK that is essential for designing chemistry assessment items. During this process, teachers considered multiple knowledge bases in addition to the "knowledge of content" knowledge base. For example, teachers communicated ideas about the "knowledge of students" knowledge base by discussing student prior knowledge (Celine: <i>If they don't remember [water's polarity] it doesn't matter...</i> ) and the student response (Celine:	



Table 6: Ashton, Emmerson, Claude item generation EO1

Statements	Assessment item
Ashton: <i>Student will be able to correctly predict what happens when a salt dissolves in water. What if we just say calcium chloride plus water with the model and a symbol? Yeah, I dunno. I'm just thinking something simple.</i>	Ca(NO <sub>3</sub> ) <sub>2</sub> (s) dissolves in water. a. What will you observe?
Claude: <i>I mean they do both in [the lab]. So correctly predict we would say symbolically and particulate but 'what happens.' Don't we have to include macroscopically as well? The expression 'what happens' that's essentially covering all three levels.</i>	b. Write the equation that describes this process. c. Draw a particulate model of the Ca(NO <sub>3</sub> ) <sub>2</sub> (aq) after it is all dissolved.
Ashton: <i>How about this, Calcium chloride, which is commonly used in the wintertime.</i>	
Claude: <i>I wouldn't use chloride because chloride is what's in [the lab].</i>	
Ashton: <i>Alright. How about calcium bromide?</i>	
Emmerson: <i>The question doesn't have to cover. It's one question for that standard.</i>	
Claude: <i>It could be three parts, right?</i>	
Emmerson: <i>I suppose.</i>	
Claude: <i>So, the question would simply be. Calcium nitrate dissolves in water. What would you observe with your eyes? Express it symbolically. Express the process symbolically.</i>	

In both vignettes, the teachers' considerations about how items. Throughout item generation teacher communicated ideas appropriately represent information impacted the processes from the knowledge bases described by the Refined Consensus Model during item generation (Hume *et al.*, 2019).

establishing an appropriate task and establishing what content assess, although the process for establishing how to represent **Research Question 1:** What is the role of high school chemistry teachers' pedagogical content knowledge when information was discussed separately of the other two processes generating planned formative assessment items for a solubility Multiple studies state the importance of representational level a means of perceiving chemical information (Johnstone, 1991; Taber, 2013). In this investigation, representational level

communicated as part of teachers' considerations for how information be represented in the assessment item as well as how information should be represented in the students' responses appropriately demonstrate competency of the task and chemistry content.

## Conclusions

When generating planned assessment items for a solubility lab, the knowledge bases embedded within high school chemistry teachers' PCK served to inform several "item generation processes." The teachers were provided with instructional materials (i.e., lab and EO) to use when generating assessment items. The results shown illustrate the need for teachers to

Table 7: Group 1 discourse during generation of assessment item for EO1

Statements	Assessment item
Celine <i>Let's have them dissolve. So, he used copper chloride so we don't want to use copper chloride. Let's just use sodium chloride. I mean honestly the kids are familiar with sodium chloride.</i>	Which choice below correctly expresses what happens to the sodium particles in NaCl when sodium chloride dissolves in H <sub>2</sub> O?
Anne <i>Right.</i>	a. Na
Celine <i>Right. He never uses sodium chloride, so let's use sodium chloride as an assessment question. When we do multiple choice test question, which choice below correctly expresses what happens to sodium chloride when it dissolves in water. Now, we can do symbolic, we can do particulate, we can do macroscopic. I say we give a macroscopic description. No, a symbolic because he's really stressing the symbolic in [the lab]. Everything he's done is supposedly symbolic, macro.</i>	b. Na <sup>+</sup> c. NaH <sub>2</sub> O d. None of the above
Anne <i>Right.</i>	
Celine <i>So, we could give them a choice of a. just the symbol Na. b. Na positive c. NaH<sub>2</sub>O because it bonds with water. Right?</i>	
Anne <i>Yeah, those are all logical choices. You don't always have to have four choices.</i>	
Celine <i>That's true.</i>	

consider both the content to be assessed in the item as well as how to construct the item to elicit student knowledge about content. The methods for how to construct the item to elicit student knowledge about certain content was informed by the “knowledge of assessment” knowledge base. This knowledge base was commonly communicated by teachers when considering how to ensure the student response required an appropriate task. Although the refined consensus model depicts each of the knowledge bases as interconnected, research showed the “knowledge of curriculum” knowledge base is often activated in conjunction with other knowledge bases (Hume *et al.*, 2019). For example, the “knowledge of curriculum” knowledge base was communicated by teachers when reasoning for how the construction of the assessment item should be connected to the content and the task being assessed in the item. Future investigations could further explore the interrelatedness between PCK knowledge bases to better understand how each is employed during various teacher tasks.

**Research Question 2:** What processes do high school chemistry teachers undergo when enacting their pedagogical content knowledge when designing planned formative assessment items for a solubility lab?

Throughout the PD activity, the PCK knowledge bases informed several “assessment item processes” that teachers underwent during assessment item generation. These processes were recognized as establishing an appropriate task to elicit student knowledge, establishing appropriate content to assess content to provide for the student, and establishing appropriate information should be represented in the item’s stem and in the student response. These processes represent the chemistry teacher’s enactment of their personal PCK. Essentially, the processes are what the teacher does to take the knowledge they have and apply it to generate a product (i.e., the assessment item). Each of the three processes are identifiable separately through teacher discourse; however, the refined consensus model of PCK indicates that these processes are likely interrelated, as are PCK knowledge bases (Hume *et al.*, 2019). As such, each of the three processes identified was determined to be essential for teachers to undergo while designing chemistry-specific assessment items.

## Limitations

The participants are experienced teachers who have been found to exhibit critical friendship after several years of participating in PD together (Schafer and Yeziarski, 2020). This camaraderie between participants serves as a strength as well as a limitation for this study. Participants’ willingness to share thoughts and ideas (and critique those of their peers) led to in-depth discourse that allowed for a fine-grained characterization of assessment beliefs and practices. However, this level of openness may not be able to be replicated with other groups of teachers, hindering generalizability. The teachers’ participation in several years of PD presents an additional limitation. The teachers have experience with student-centered methods, limiting generalizability of the findings to those who would not have a lab such as *What’s the Solution?* in their classrooms.

Another limitation is the likelihood that all ideas were not openly communicated. Although these teachers have been shown to willingly share ideas in the past, they likely did not share all their thoughts and ideas leading to a possibly incomplete characterization of their enacted PCK. To this end, it is important to recognize that the teachers were given ample opportunity to generate assessment items. Teachers ended the item generation period of their own volition; the items developed were perceived as best-possible quality by the teachers before critique from peers. We would like to recognize our support of these individuals (and all teachers) and hope these findings lead to collective growth.

A further limitation of this work is that the analysis focused on the ideas used to generate assessment items for a specific chemistry topic and not the quality of the assessment items generated or the content accuracy of teacher ideas. As such, the final items presented in this study are not necessarily examples of high-quality items. Further investigation is required to understand how the quality of any individual knowledge base (such as a teacher’s content knowledge) influences teacher enactment of PCK.

## Implications for Research and Future Work

Here evidence of the knowledge bases that inform and are informed by high school chemistry teacher PCK about solubility lead to implications about the role of PCK on assessment item development. The findings presented imply that a teacher’s PCK has an observable influence on the generation of planned formative assessment items when characterized by its embedded knowledge bases. When generating assessment items, a teacher likely undergoes particular processes for translating personal-PCK into enacted-PCK (establishing an appropriate task for the student to accomplish, establishing the content to assess versus what content is provided to the student, establishing an appropriate representation of the content in the stem and in the student response). However, further investigation is needed to understand if these processes are consistent from topic to topic and which factors contribute to teacher ability to enact their personal-PCK. Understanding how PCK is enacted by teachers and characterizing how a teacher’s PCK influences the products they develop can lead to more precise models of teacher knowledge as well as improved support for high school chemistry teachers implementing reformed practices in their classrooms. Future studies can further investigate the relationship between these processes and how teacher use of individual knowledge bases contributes to the ability to engage in these processes.

## Implications for Teaching and Future Work

Results from this investigation indicate that a teacher likely undergoes specific, identifiable processes to enact their PCK during assessment item development. Undergoing these processes in some way contributed to the development of formative assessment items for the teachers in this study.

Although the quality of the assessment items was not investigated, teachers undergoing these processes likely have measurable characteristics (e.g., chemistry content knowledge) that would contribute to assessment item quality. This implies that chemistry teachers may benefit from considering how they enact PCK during assessment item design. Teachers may scaffold their considerations through the lens of the knowledge bases embedded within PCK, as was done in this study. Future investigations could characterize the how the processes from the study herein influence assessment item quality, how specific PCK knowledge bases afford (or hinder) teacher ability to employ the processes, and how a teacher's ability to carry out the processes affects assessment item quality.

Similar investigations have found that as teacher PCK strengthens, so too does their ability to foster student understanding (Pajares, 1992; Marzabal *et al.*, 2018). A similar trend may be observable for the relationship between PCK and a teacher's ability to generate high-quality assessment items. High school chemistry teachers would likely benefit by reflecting on their assessment item development processes. Again, teachers may scaffold their reflections through the lens of the individual knowledge bases. Taking the time to consider how knowledge is enacted within a particular chemistry context could both bolster productive skills and highlight gaps in assessment design practices that influence assessment design. As such, professional development designers should provide chemistry teachers with opportunities for sustained professional development that explicitly connects the process of assessment item design to other knowledge bases as teachers transform their knowledge into classroom tools and tasks for assessment.

## APPENDIX

### Appendix A

Knowledge Base (Knowledge of...)	Ideas	Description (Statements that directly communicate the ...)
Students	Student Prior Knowledge	student prior experiences aligned to the content, task, or representational level of the item
	Student Response	student ability to respond to the item or potential student response to the item
Content	Representational Level	Johnstone's level emphasized in the stem or student response
	Dissolving	phenomenon of a substance dissolving in solution
	Substance	atoms, ions, particles, molecules, compounds involved in the phenomenon
	Electrostatic Interaction	strength or presence of attractive or repulsive forces due to electric charge (or partial electric charge)
Assessment	Item Format	arrangement of the stem or student response
Curriculum	Learning Activity	<i>What's the Solution?</i> inquiry activity
	Educational Objective	educational objectives provided for the <i>What's the Solution?</i> inquiry activity
	State/National Standards	state and or national standards for chemistry

## 765 AUTHOR INFORMATION

766 Corresponding Author

767 \*E-mail: [yeziere@miamioh.edu](mailto:yeziere@miamioh.edu)

## 768 Conflicts of interest

769 There are no conflicts to declare.

## 770 Acknowledgments

771 We thank the high school chemistry teachers for participating  
772 in this project. We also thank the Yezierski and Bretz research  
773 groups at Miami University for their feedback and guidance.  
774 This material is based upon work supported by the U. S. National  
775 Science Foundation under Grant No. (DRL-1118749).

## 776 References

- 777 1. Abell S. K., (2008), Twenty Years Later: Does pedagogical  
778 content knowledge remain a useful idea? *Int. J. Sci. Educ.*,  
779 **30**(10), 1405–1416.

- 1  
2  
3 780 2. Abell S. K. and Siegel M. A., (2011), Assessment Literacy: What  
4 781 Science Teachers Need To Know and Be Able To Do. in  
5 782 Corrigan D., Dillon J., and Gunstone R. (eds.), *The*  
6 783 *Professional Knowledge Base of Science Teaching*. New  
7 784 York, NY: Springer, pp. 205–221. 841
- 8 785 3. Bell B. and Cowie B., (2001), The Characteristics of Formative  
9 786 Assessment in Science Education. *Sci. Educ.*, **85**(5), 536–543  
10 787 4. Black P. and Wiliam D., (1998), Inside the Black Box: Raising  
11 788 Standards Through Classroom Assessment. *Phi Delta*  
12 789 *Kappan*, **80**(2), 139–148. 846
- 13 790 5. Buck G. A. and Trauth-Nare A. E., (2009), Preparing teachers  
14 791 to make the formative assessment process integral to science  
15 792 teaching and learning. *J. Sci. Teacher Educ.*, **20**(5), 475–494  
16 793 6. Cisterna D. and Gotwals A. W., (2018), Enactment of Ongoing  
17 794 Formative Assessment: Challenges and Opportunities for  
18 795 Professional Development and Practice. *J. Sci. Teacher Educ.*  
19 796 **29**(3), 200–222. 853
- 20 797 7. Clinchot M., Ngai C., Huie R., Talanquer V., Banks G., Weintraub  
21 798 M., et al., (2017), Better Formative Assessment: Making  
22 799 formative assessment more responsive to student needs. *Sci*  
23 800 *Teach.*, **84**(3), 69–75. 857
- 24 801 8. Coffey J. E., Hammer D., Levin D. M., and Grant T., (2011), The  
25 802 missing disciplinary substance of formative assessment. *J.*  
26 803 *Res. Sci. Teach.*, **48**(10), 1109–1136. 860
- 27 804 9. DeLuca C., Valiquette A., Coombs A., LaPointe-McEwan D.,  
28 805 Luhanga U., (2018), Teachers' Approaches to Classroom  
29 806 Assessment: A Large-Scale Survey. *Assess. Educ. Princ.*  
30 807 *Policy Pract.*, **25**(4), 355–375. 864
- 31 808 10. Dini V., Sevian H., Caushi K., and Orduña Picón R., (2020),  
32 809 Characterizing the formative assessment enactment of  
33 810 experienced science teachers. *Sci. Educ.*, **104**(2), 290–325. 867
- 34 811 11. Dubey P. and Geanakoplos J., (2010), Grading exams:  
35 812 100,99,98,... or A,B,C? *Games Econ. Behav.*, **69**(1), 72–94. 869
- 36 813 12. Edwards F., (2013), Quality assessment by science teachers:  
37 814 Five focus areas. *Sci. Educ. Int.*, **24**(2), 212–226. 871
- 38 815 13. Furtak E. M., Kiemer K., Circi R. K., Swanson R., de León V.,  
39 816 Morrison D., and Heredia S. C., (2016), Teachers' formative  
40 817 assessment abilities and their relationship to student learning  
41 818 findings from a four-year intervention study. *Instr. Sci.*, **44**(3),  
42 819 267–291. 876
- 43 820 14. Gess-Newsome J., (2015), A model of teacher professional  
44 821 knowledge and skill including PCK: Results of the thinking  
45 822 from the PCK Summit. in Berry A., Friedrichsen P. M., and  
46 823 Loughran J. (eds.), *Re-examining Pedagogical Content*  
47 824 *Knowledge in Science Education*. New York, NY: Routledge,  
48 825 pp. 28–43. 882
- 49 826 15. Gómez M. del C. and Jakobsson A., (2014), Everyday classroom  
50 827 assessment practices in science classrooms in Sweden. *Cult*  
51 828 *Stud. Sci. Educ.*, **9**(4), 825–853. 885
- 52 829 16. Harris C. J., Krajcik J. S., Pellegrino J. W., Mcelhaney K. W.,  
53 830 DeBarger A. H., Dahsah C., et al., (2016), Constructing  
54 831 Assessment Tasks that Blend Disciplinary Core Ideas,  
55 832 Crosscutting Concepts, and Science Practices for Classroom  
56 833 Formative Applications Center for Technology in Learning  
57 834 Menlo Park, CA: SRI International. 891
- 58 835 17. Harshman J. and Yeziarski E., (2017), Assessment Data-driven  
59 836 Inquiry: A Review of How to Use Assessment. *Sci. Educ.*,  
60 892
18. Henderson C., Yerushalmi E., Kuo V. H., Heller P., and Heller  
K., (2004), Grading student problem solutions: The challenge  
of sending a consistent message. *Am. J. Phys.*, **72**(2), 164–  
169.
19. Herman J. L., Webb N. M., and Zuniga S. A., (2005),  
Measurement Issues in the Alignment of Standards and  
Assessments: A Case Study. CSE Report 653. *Natl. Cent. Res.*  
*Eval. Stand. Student Test.*, **20**(1), 101–126.
20. Hoffman C. K. and Medsker K. L., (1983), Instructional  
analysis: The missing link between task analysis and  
objectives. *J. Instr. Dev.*, **6**(4).
21. Hume A., Cooper R., and Borowski A., (2019), Repositioning  
pedagogical content knowledge in teachers' knowledge for  
teaching science.
22. Irons A., (2008), *Enhancing Learning Through Formative  
Assessment and Feedback*, New York, NY: Routledge.
23. Johnstone A. H., (1991), Why is science difficult to learn?  
Things are seldom what they seem. *J. Comput. Assist. Learn.*,  
**7**(2), 75–83.
24. Jonassen D. H., Tessmer M., and Hannum W. H., (1999), *Task  
Analysis Methods for Instructional Design*, Mahwah, New  
Jersey: Lawrence Erlbaum Associates.
25. Magnusson S., Krajcik J., and Borko H., (1999), Nature, sources,  
and development of pedagogical content knowledge for  
science teaching. *Examining Pedagog. content Knowl.*, 95–  
132.
26. Martone A. and Sireci S. G., (2009), Evaluating Alignment  
Between Curriculum, Assessment, and Instruction. *Rev. Educ.*  
*Res.*, **79**(4), 1332–1361.
27. Marzabal A., Delgado V., Moreira P., Barrientos L., and Moreno  
J., (2018), Pedagogical Content Knowledge of Chemical  
Kinetics: Experiment Selection Criteria to Address Students'  
Intuitive Conceptions. *J. Chem. Educ.*, **95**(8), 1245–1249.
28. Maxwell J. A., (2013), *Qualitative Research Design: An  
Interactive Approach*, 3rd ed. Knight V. (ed.) Thousand Oaks,  
California: SAGE Publications.
29. McDonald F. J., (1964), Meaningful Learning and Retention:  
Task and Method Variables. *Rev. Educ. Res.*, **34**, 530–544.
30. Merrill M. D., (2007), A Task-Centered Instructional Strategy. *J.*  
*Res. Technol. Educ.*, **40**(1), 5–22.
31. National Academies of Sciences, Engineering and M., (2015),  
*Science Teachers' Learning: Enhancing Opportunities,  
Creating Supportive Contexts*, Washington, DC: National  
Academies Press.
32. National Research Council, (2014), *Developing Assessments for  
the Next Generation Science Standards*, Washington, DC: The  
National Academies Press.
33. NGSS Lead States, (2013), *Next Generation Science Standards:  
For States, by States (Appendix F – Science and Engineering  
Practices)*. *Achieve, Inc. behalf twenty-six states partners that  
Collab. NGSS*, (November), 1–103.
34. Opre D., (2015), Teachers' Conceptions of Assessment.  
*Procedia - Soc. Behav. Sci.*, **209**, 229–233.
35. Pajares M. F., (1992), Teachers' Beliefs and Educational  
Research: Cleaning Up a Messy Construct. *Rev. Educ. Res.*,  
**62**(3), 307–332.

36. Park S. and Chen Y. C., (2012), Mapping out the integration of the components of pedagogical content knowledge (PCK): Examples from high school biology classrooms. *J. Res. Sci. Teach.*, **49**(7), 922–941.
37. Park S. and Oliver J. S., (2008), Revisiting the conceptualisation of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Res. Sci. Educ.*, **38**(3), 261–284.
38. Polikoff M. S. and Porter A. C., (2014), Instructional Alignment as a Measure of Teaching Quality. *Educ. Eval. Policy Anal.*, **36**(4), 399–416.
39. Remesal A., (2011), Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teach. Teach. Educ.*, **27**(2), 472–482.
40. Ruiz-Primo M. A. and Furtak E., (2007), Exploring Teachers' Informal Formative Assessment Practices and Students' Understanding in the Context of Scientific Inquiry. *J. Res. Teach.*, **44**(1), 57–84.
41. Ruiz-Primo M. A., Li M., Wills K., Giamellaro M., Lan M. C., Mason H., and Sands D., (2012), Developing and Evaluating Instructionally Sensitive Assessments in Science. *J. Res. Sci. Teach.*, **49**(6), 691–712.
42. Sandlin B., Harshman J., and Yeziarski E., (2015), Formative Assessment in High School Chemistry Teaching: Investigating the Alignment of Teachers' Goals with Their Items. *J. Chem. Educ.*, **92**(10), 1619–1625.
43. Schafer A. G. L. and Yeziarski E. J., (2020), Chemistry critical friendships: Investigating for chemistry-specific discourse within a domain-general discussion of best practices for inquiry assessments. *Chem. Educ. Res. Pr.*, **21**(1), 452–468.
44. Shepard L. A., (2000), The Role of Assessment in a Learning Culture. *Educ. Res.*, **29**(7), 4–14.
45. Shulman L., (1987), Knowledge and Teaching: Foundations of the New Reform. *Harv. Educ. Rev.*, **57**(1), 1–21.
46. Shulman L. S., (1986), Those Who Understand: Knowledge Growth in Teaching. *Educ. Res.*, **15**(2), 4–14.
47. Siegel M. A., Cite S., Muslu N., Murakami C. D., Burcks S. M., Izci K., and Nguyen P. D., (2019), Attending to assessment problems of practice during community-centered professional development. *Int. J. Educ. Res.*, **95**, 190–199.
48. Suskie L., (2009), *Assessing Student Learning: A Common Sense Guide*, 2nd ed. San Fransisco, CA: Jossey-Bass.
49. Taber K. S., (2013), Revisiting the chemistry triplet: Drawing upon the nature of chemical knowledge and the psychology of learning to inform chemistry education. *Chem. Educ. Res. Pract.*, **14**(2), 156–168.
50. Tamir P., (1991), Professional and personal knowledge of teachers and teacher educators. *Teach. Teach. Educ.*, **7**(3), 263–268.
51. Target Inquiry at Miami University.
52. Toledo S. and Dubas J. M., (2017), A Learner-Centered Grading Method Focused on Reaching Proficiency with Course Learning Outcomes. *J. Chem. Educ.*, **94**(8), 1043–1050.
53. Tomanek D., Talanquer V., and Novodvorsky I., (2008), What Do Science Teachers Consider When Selecting Formative Assessment Tasks? *J Res Sci Teach.*, **45**(10), 1113–1130.
54. Towndrow P. A., Tan A.-L., Yung B. H. W., and Cohen L., (2010), Science Teachers' Professional Development and Changes in Science Practical Assessment Practices: What are the Issues? *Res. Sci. Educ.*, **40**(2), 117–132.
55. Towns M. H., (2014), Guide to developing high-quality, reliable, and valid multiple-choice assessments. *J. Chem. Educ.*, **91**(9), 1426–1431.
56. Underwood S. M., Posey L. A., Herrington D. G., Carmel J. H., and Cooper M. M., (2018), Adapting Assessment Tasks to Support Three-Dimensional Learning. *J. Chem. Educ.*, **95**(2), 207–217.
57. Wolf D., Bixby J., Glenn J., and Gardner H., (1991), To Use Their Minds Well: Investigating New Forms of Student Assessment. *Rev. Res. Educ.*, **17**(1), 31–74.
58. Yan Z. and Cheng E. C. K., (2015), Primary teachers' attitudes, intentions and practices regarding formative assessment. *Teach. Teach. Educ.*, **45**, 128–136.
59. Yeziarski E. J. and Herrington D. G., (2011), Improving practice with target inquiry: high school chemistry teacher professional development that works. *Chem. Educ. Res. Pr.*, **12**(3), 344–354.