Chemistry Education Research and Practice

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Chem. Educ. Res. Pract.*, 2023, **24**, 914

Assessor in action: assessment literacy development in a biochemistry context

Stephanie Feola,[a] Paula P. Lemons,[b] Jennifer A. Loertscher,[c] Vicky Minderhout[c] and Jennifer E. Lewis[a]

Received 20th December 2022,
Accepted 5th March 2023

DOI: 10.1039/d2rp00334a

rsc.li/cerp

Instructors make assessment decisions based on their knowledge and experiences. Assessment practice is an essential element of instruction, and the outcomes of assessments have a broad impact on both students and instructors. Efforts to provide strengths-focused, relevant professional development support regarding assessment are enhanced by greater understanding of the complex nature of assessment practices. In this study, the Teacher Assessment Literacy in Practice (TALiP) framework was used to guide our investigation of one biochemistry instructor's assessment literacy, relevant to her integration of a biochemistry threshold concept, the physical basis of interactions (PBI), into her course. Qualitative framework analysis was used to examine classroom artifacts and interview data to reveal that community support and self-reflection influenced the instructor's enactment of specific assessments aligned with her instructional goals. Additionally, the instructor was seen to leverage assets to develop her assessment literacy within a single semester.

Introduction

Supporting the adoption and enactment of effective teaching practices in the classroom is a major focus of faculty professional development in higher education (Gillespie *et al.*, 2010). Faculty knowledge of assessment techniques and their interpretation of assessment outcomes are significant points of focus for professional development offered by centers for teaching and learning (CTL) (Beach *et al.*, 2016). As such, there has been a call for faculty's assessment literacy to be a focus of sustained professional development and support (Massey *et al.*, 2020). Assessments are tools used by instructors to determine whether students are achieving course benchmarks (Luxford and Holme, 2016; Tienson-Tseng, 2019) and allow instructors and researchers insight into student knowledge and skills. Increasingly, instructors engage with assessment as part of departmental efforts to track students' performance, yet many instructors report challenges (Emenike *et al.*, 2013a). Researchers often consider how instructors' assessment choices and interpretations affect students, but these choices also impact the instructors themselves (Offerdahl and Tomanek, 2011; Offerdahl and Montplaisir, 2014). For example, assessments can influence an instructor's knowledge of students' understanding and of pedagogical techniques (Sadler, 2012; Talanquer *et al.*, 2015; Demirdogen and Korkut, 2021; Herridge and Talanquer, 2021). Previous work has shown that chemistry instructors have knowledge of terminology related to assessment and knowledge about tools of summative and formative assessment, however this knowledge is distributed differently depending on institution type, and teaching and research experience (Emenike *et al.*, 2013b; Raker *et al.*, 2013; Raker and Holme, 2014; Coombs *et al.*, 2018; Herridge and Talanquer, 2021; Gibbons *et al.*, 2022). Differences in beliefs on the purpose of assessment also differ depending on whether the chemistry instructor has a science or education background (Demirdogen and Korkut, 2021).

Assessment development requires that instructors carefully consider how to craft assessment questions that elucidate student knowledge, how to interpret the information from those assessments, and how the assessment outcomes will impact students. Schafer *et al.* (2021) examined how instructor intent, learning objectives, and assessment items are aligned with one another. They found that there was general alignment however there was frequently a disconnect between assessment tasks and the instructor intention when designing the assessment tied to their course objectives. The misalignment was not necessarily due to a deficit of the instructor but, rather, to the differences in their expectations and their chosen assessment approach. Additionally, Schafer and Yezierski (2021a) have indicated that there is a need to examine the link between what assessment items ask students to do and what tasks

[a] *Department of Chemistry and Center for Improvement of Teaching and Research on Undergraduate STEM, University of South Florida, Tampa, FL 33620, USA. E-mail: sfeola@usf.edu, jennifer@usf.edu*
[b] *Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA, 30602, USA. E-mail: plemons@uga.edu*
[c] *Department of Chemistry, Seattle University, Seattle, WA, 9812, USA. E-mail: loertscher@seattleu.edu, vicky@seattleu.edu*

student's perform which may help instructors design assessments that align with their assessment intent. Schafer and Yezierski (2021b) also indicate that examining the processes by which individual instructors translate knowledge of assessment into assessment practice will allow for more insight into assessment literacy. By extension, to create relevant professional development opportunities, it is important to examine the assessment assets that instructors bring and enact in their teaching practice in order to create relevant professional development opportunities. Instructor assets include analytical skills, content knowledge, socio-cultural knowledge, and experiences (Mansour, 2009; Harper, 2010; Samuelson and Litzler, 2016; Cowie and Trevethan, 2021; Abell and Sevian, 2021; DeLuca *et al.*, 2021).

The TALiP model provides insight into assessment literacy and assessor identity

Assessment literacy provides a lens to examine instructor assets related to their assessment knowledge base and how those assets can be leveraged to promote student learning. All instructors bring disciplinary expertise as an asset, and over time they gain experience applying this knowledge to developing and interpreting assessments (Scarino, 2013; Cowie *et al.*, 2014; Coombs *et al.*, 2018). The knowledge of how concepts within a discipline are interrelated is an assessment asset that has been well-characterized in biochemistry education contexts (Tansey *et al.*, 2013; White *et al.*, 2013; Loertscher *et al.*, 2014; Bell *et al.*, 2019). Yet, assessment literacy encompasses not only disciplinary knowledge and views of assessment, but also involves understanding the rationale of assessment, implementing assessment schemes, and reflecting on outcomes of assessments. The instructor is at the heart of assessment literacy. Therefore, if we understand an individual instructor's process of translating assessment knowledge into assessment practice (Schafer and Yezierski, 2021a; Schafer *et al.*, 2021), we can examine the elements at play in assessment literacy. The Teacher Assessment Literacy in Practice (TALiP) framework (Xu and Brown, 2016, original framework in Appendix 2) can be used to examine an instructor's assessment practice through a multi-faceted lens. The TALiP framework was developed through a meta-analysis of assessment literacy literature. It acknowledges the multiplicity of factors that are at play when an instructor makes assessment choices, with how they perceive themselves as an assessor as the apex of a pyramid of factors. This framework informs our conceptualization and understanding of the assets at work in assessment literacy.

The instructor-as-assessor identity is complex and dynamic

The peak element of the TALiP framework is an instructor's identity as an assessor, which is built through assessment experience and synergistic interaction of all elements and levels of the framework (The full original framework has been provided in the Appendix 2). Due to the centrality of assessment in chemistry and biochemistry education, it is important to understand how instructors build their assessor identities and perceive the function of assessment in their classroom (McMillan, 2003; Looney *et al.*, 2018; Coombs *et al.*, 2020, Demirdogen and Korkut, 2021). Specifically, assessor identity is built through learning from assessment outcomes and reflecting on those outcomes to connect them to an instructor's own assessment practices (DeLuca *et al.*, 2018 and 2019). An instructor's actions within the classroom both are informed by and create a context for their assessment decisions. These decisions include not only which concepts to assess and which modes of assessment to use, but also how to analyze and interpret responses to assessment items (Allal, 2013; Demirdogen and Korkut, 2021; Herridge and Talanquer, 2021). This assessment sense-making activity, essential to an instructor's self-reflective cycles, relies upon the interactions of all of the elements of the TALiP framework and fosters the development of assessor identity.

Instructors have varying experiences with assessment training, which can be formal or informal (Izci and Siegel, 2019; Schultz *et al.*, 2018). Instructors' experiences of assessment training are likely influenced by their experiences of assessment as a student, their instructional experiences, talking with their peers, or formal professional development. Exposure to different assessment styles and techniques can also increase an instructor's self-awareness and self-efficacy regarding the utility of assessment and its place in their teaching practice (Offerdahl and Tomanek, 2011; Scarino, 2013; Levy-Vered and Nasser-Abu Alhija, 2018; Coombs *et al.*, 2020). Offerdahl and Tomanek (2011) have shown that a team of instructors who implemented new formative assessment techniques in their classrooms gained insight into student thinking and that the instructors' conceptions about usefulness of assessment changed over time. Further, Herridge *et al.* (2021) has indicated that examination of an instructor's personal philosophies regarding assessment is necessary to understand their enacted assessment practices and perceptions of assessment outcomes. The TALiP framework is aligned with this perspective.

Community is important for the development of assessor identity

Participation in a professional development environment can influence assessment literacy in practice and assessor identity, especially by leveraging interaction with peers. Specific professional development has been used to impact instructors' assessment skills and techniques (Lukin *et al.*, 2004; Frey and Fisher, 2009; Koh, 2011; DeLuca and Klinger, 2010; DeLuca and Johnson, 2017; Yan *et al.*, 2018). An instructor's assessment literacy can be influenced through interaction with colleagues focused on discipline-specific ideas and how to assess them (Adie, 2013; Adie *et al.*, 2017; Allal, 2013; Fulmer *et al.*, 2015). These interactions are different from one-off professional development opportunities in that they involve the development of a community that can be an ongoing resource for the instructor (Adie, 2013; Levy-Vered and Nasser-Abu Alhija, 2018; Sadler and Reimann, 2017; Herppich *et al.*, 2018). Such communities are called communities of practice.

A community of practice (CoP) is a group of people who come together in pursuit of a common interest, such as sharing experience to solve a problem, creating common resources, generating new knowledge together, and/or communicating knowledge outside of the community (Wenger, 1999). There have been indications that participation in professional development activities can break down barriers of isolation and allow for the development of a community (Hadar and Brody, 2010). Communities of practice are self-motivated and, through the process of working together, create their own standards of practice (Gehrke and Kezar, 2017). Motivation to contribute usually comes with a reward that the individual member receives from the community; it could be knowledge, camaraderie, sense of accomplishment, or an increase in self-efficacy (Takahashi, 2011; Austin and Murray, 2019), which are important for the development of an assessor identity (Levy-Vered and Nasser-Abu Alhija, 2018). Therefore, it is important to give instructors the opportunity to interact with a community of peers to enrich their assessment literacy and grow as assessors (Lukin *et al.*, 2004; Wyatt-Smith *et al.*, 2010; Forsberg and Wermke, 2012). CoPs have been useful to support faculty professional development broadly in STEM fields (Tomkin *et al.*, 2019) and in specific subdisciplines including inorganic chemistry (Watson *et al.*, 2020) and biochemistry (Loertscher, 2011).

Assessment literacy leverages an instructor's assets

A holistic consideration of instructors' ability and knowledge related to assessment is not fully reflected in education research publications. Current literature on assessment literacy in chemistry is often interpreted and reported using a deficit approach. Although some studies seek to understand instructor perspectives and knowledge in relation to teaching and assessment, many studies report on assessment literacy using one of three deficit frames. First, focus on instructor gaps of knowledge or lack of awareness that highlight perceived deficiencies of instructors with phrases like "*quality and productivity of some specific interpretations was not high due to the teachers' own misunderstandings or lack of awareness of naïve patterns of reasoning*" (Talanquer *et al.*, 2015). Second there is a value judgment of instructors lacking depth or progress in their assessment/teaching which can be seen in phrases like "what c*ompounds the problem of depth versus breadth is instructors' understanding of how students learn. A relative lack of familiarity often leads to the failure to use instructional strategies [...]*" (Paek and Holme, 2013). Third, there is a tendency to reinforce a hierarchy of novice (or naïve) to expert which reinforces the idea that expert skill is the only valid way of knowing, seen in this quote: "*lack of teaching experience limited her to elicit and address common misconceptions and use assessment results*" (Izci and Siegel, 2019). An asset-based lens would allow for a way to reconceptualize how we view an instructor's assessment literacy in practice.

Assessment literacy is a multi-faceted concept that integrates assessment assets, including disciplinary expertise and knowledge of assessment practices, with affective and cognitive dimensions of assessment (Xu and Brown, 2016). Assessment literacy is dynamic and can be strengthened over time through thoughtful consideration of what questions we as instructors ask students and for what purpose (Crisp, 2012). Previous research has posited that there is a "spectrum of assessment thinking" that influences when an instructor chooses to implement a change in their assessment and teaching practice (Offerdahl and Tomanek, 2011 p. 791). With assessments having high stakes for students, assessment implementation can have implications on equity of outcomes for students (Montenegro and Jankowski, 2017 and 2020). It is important to remember that while there is discussion about shifting to an asset-based frame for students, especially minoritized students, many of those same students become instructors and the asset view should not end when they take on the additional identity element of instructor and assessor. Viewing instructors as learners, provides us with the opportunity to explore how they engage with assessment in a way that is relevant to their instructional and institutional context. Changing to an asset-based lens reframes instructors' assessment practice as their utilization of strengths and how that can translate into their identity as assessor. It also allows us to examine the complexity of the knowledge assets they bring into their decision-making regarding assessment and how that informs their identity.

There are multiple asset-based approaches to education and pedagogy including, but not limited to, culturally relevant pedagogies (Ladson-Billings, 1995; Ball and Ladson-Billings, 2020), culturally sustaining/revitalizing pedagogies (Jacob *et al.*, 2018; McCarty and Lee, 2014) funds of knowledge (Moll *et al.*, 1992; González *et al.*, 2006) and "*funds of identity*" (Esteban-Guitart and Moll, 2014). While this study does not ascribe to a specific asset-based approach, we are using an asset-based frame in our approach to identify assessment literacy assets. Examining instructor assets altered how we framed our evaluative view of what knowledge instructors decide to leverage while enacting and interpreting assessments. Examining assessment literacy in practice can provide us with a way to connect instructor assets and their assessor identity. Therefore, we used the TALiP to specifically tie assets to elements of assessment literacy in the framework inclusive of assessor identity.

Research questions

In this study we explore the assessment literacy in practice of an instructor who, at the time, had recently shifted from teaching biology to biochemistry. Her shift provided an opportunity to look at how her assessment literacy in practice was influenced by her membership in a biochemistry CoP. Biochemistry by nature is an interdisciplinary field and is approached by both students and instructors from multiple backgrounds including, but not limited to, biology, chemistry, biophysics, and pharmacology. As a natural result, instructors make pedagogical decisions based on experience and disciplinary knowledge. The more insight into the complex nature of instructional choices and perspectives related to assessment in biochemistry, the better we can develop strengths-based and relevant professional development and instructional support for biochemistry instructors. This work seeks to further understand how the components of instructor assessment assets and community all play a part in their assessment identity in a biochemistry context. We focused the following research questions:

(1) How was the instructor's assessment literacy in practice influenced by membership in a biochemistry community of practice?

(2) What assets did the instructor leverage to assess student understanding in a biochemistry context?

(3) In what ways did the instructor's learning from and reflection on assessment outcomes influence her identity as an assessor?

Methodology

Naturalistic inquiry and trustworthiness

Multiple paradigms are available to researchers seeking to examine and understand phenomena. Every paradigm has its own ontology, epistemology, and axiology along with traditions regarding generalizability and causal linkages. This research study was guided by naturalistic inquiry (Lincoln and Guba, 1985). Researchers operating in this paradigm have found it important to address questions of trustworthiness explicitly. Specifically, the naturalistic inquiry perspective has four major criteria for trustworthiness (Guba and Lincoln, 1982; Lincoln and Guba, 1985 pg. 300–327): credibility, transferability, dependability, and confirmability.

First, it is important to note that naturalistic inquiry seeks transferability rather than generalizability. The transferability criterion for this case study is addressed through the thick description of the case study, its context, and the interpretations of the data, which are included in this article and the Appendix. Regarding credibility, dependability, and confirmability, multiple data sources, prolonged engagement, and detailed bookkeeping regarding analysis decisions support these elements of trustworthiness. This research involved long-term member checks with the research participant who was the focus of inquiry. This high degree of interaction is recognized in the authorship structure of this paper. Our data set consists of multiple, overlapping sources: multiple interviews, classroom artifacts including assessments and teaching materials, a grant proposal involving assessment written by the participant, and the participant's personal notes and other artifacts from the biochemistry CoP's workshops. These overlapping sources supported triangulation of data analysis. Additionally, two members of the research team (VM and JAL) had prolonged engagement with the participant and the biochemistry CoP through their actions as coordinators of the biochemistry CoP (more details below). Overall, the research team has had prolonged engagement with biochemistry and higher education contexts. Both types of prolonged engagement allowed us to make sense of the materials for analysis and interpretation. Peer debriefing of data and analysis was conducted with other chemistry education researchers who were not involved with the project and context.

Context of the inquiry focus

Biochemistry CoP context. The biochemistry community of practice work was part of a National Science Foundation funded project to improve student understanding of threshold concepts in biochemistry. Threshold concepts are ideas which, once understood, lead to deep and integrated learning within a discipline (Meyer and Land, 2006). Efforts to identify threshold concepts involve disciplinary experts, students, educators, and researchers working together as a community. The primary purpose for identifying threshold concepts is to provide a starting point for focused curricular redesign, because an intentional approach to teaching a specific set of key concepts is likely to result in the greatest improvement in student learning (Entwistle, 2008; Perkins, 2008). Five threshold concepts for undergraduate biochemistry were identified by this CoP: steady state, biochemical pathway dynamics and regulation, the physical basis of interactions (PBI), thermodynamics of macromolecular structure formation, and free energy (Loertscher *et al.*, 2014; Green *et al.*, 2017).

Since 2007, this and another NSF funded project have worked to cultivate a community of over 100 biology, chemistry, and biochemistry college and university educators who are engaged in improving teaching and learning in the molecular life sciences (Murray *et al.*, 2011; Villafañe *et al.*, 2021). From this community, a total of 45 different people attended one or more three-day workshops held every summer from 2013–2016 to develop instructional and assessment materials related to threshold concepts. These workshops were known as the core collaborators workshops, and they constitute what will be referred to as the biochemistry CoP from here on. The instructor in this study, PPL, attended all four of the in-person workshops. Workshop participants included 40 faculty members, three graduate students, and two postdoctoral fellows, drawn from diverse institution types including two-year colleges, small four-year colleges, master's-granting universities, and research-universities. Workshop participants listened to presentations about advances in discipline-based education research (DBER), shared their own instructional innovations, and participated in activities to develop assessment and instructional materials related to threshold concepts. Two of the authors (VM and JAL) were co-PIs on the NSF grant that funded the workshops and as such designed and facilitated all the workshops. VM and JAL also maintained contact *via* email, video conference, and in-person meetings with biochemistry CoP members in the time between workshops.

Workshops adopted a collaborative approach with frequent opportunities for reflection. Participants worked in structured small groups to accomplish tasks related to developing and refining instructional and assessment materials. Small group activities were typically followed by large group discussion in which ideas and insights were shared. Group composition was changed several times throughout each workshop, allowing all workshop participants to work with each other to foster relationships and build community. Unstructured time was intentionally built into the workshops to enable participants to engage with each other more deeply on topics of mutual interest. Throughout the workshops, participants were asked to reflect on activities and to identify strengths, areas for further development, and insights related to either the process used to generate materials or related to the materials themselves.

Instructional context. Assessment data for this study was generated in Spring 2016 at a large, research-intensive institution in the southeastern United States. During that semester, co-author Paula P. Lemons (PPL) was teaching introductory biochemistry, a one-semester course for science and engineering students who are not biochemistry majors. Most students who take this course are preparing to pursue graduate or professional studies in health-related careers. The course covers topics including proteins, enzyme-catalyzed reactions, metabolism, and nucleic acid structure and function. Prerequisites include introductory biology, general chemistry, and the first semester of organic chemistry. In Spring 2016, introductory biochemistry enrolled 199 students. Coursework incorporated in-class case studies that required students to complete related preparatory and follow-up assignments outside of class as well. During class time, peer learning assistants who had successfully completed the course in a prior term supported student work on the case studies. Formal course assessments included four exams with multiple-choice and short-answer items. While the course was co-taught with one other instructor, the structure of the co-teaching arrangement resulted in the topics pertinent to this paper being taught and assessed by PPL only.

Data corpus

Extant data was provided by PPL. This data includes notes taken by PPL during the first biochemistry CoP workshop in the summer of 2013, a definition handout from the biochemistry CoP that PPL used for biochemistry threshold concepts reference, PPL's syllabi from three consecutive spring semesters of introductory biochemistry (2014–2016), a grant proposal regarding assessment submitted by PPL in 2013, two case studies used as in-class activities in Spring 2016, and the Spring 2016 course exams with their corresponding keys. The extant data was examined for connections to a specific threshold concept, physical basis of interactions (discussed below in Landmark 1), from which we structured our next collection of data: three in-depth semi-structured interviews with PPL in the Spring of 2018 *via* Zoom.

Interview data was collected by SF. Interview data was collected in accordance with IRB requirements for social science research, including participant consent. Each interview focused on a different aspect emergent from the extant data provided by PPL. Each set of interview questions can be found in Appendix 1. The first interview (65 min) included questions about the relationships among learning objectives, threshold concepts, and assessments. The second interview (72 min) focused on relationships between assessment literacy and enacted pedagogy, drawing on the content representation questions from Schultz *et al.*'s tertiary teaching practice CoRe questionnaire (2018). These questions were modified to align with two case study assignments developed by PPL and used in Spring 2016, such that the 'Big Ideas' were centered on examining her specific classroom activities. The third interview (55 min) focused on the instructor's membership in CoPs, including the biochemistry CoP where the threshold concepts were developed. The content of the interviews prompted additional extant data to be requested, collected, and used in the examination of the context of PPL and to help inform the data analysis.

Framework analysis

Framework analysis is a form of thematic analysis. The process of performing a framework analysis involves data familiarization, identifying a thematic framework from the data and the literature, indexing all study data within the thematic framework, charting/visualizing to summarize the indexed data, and mapping and interpretation of patterns (Ritchie and Spencer, 2002; Parkinson *et al.*, 2016; Kiernan and Hill, 2018; Goldsmith, 2021). Traditionally framework analysis results in an entirely new framework; however, for this study the TALiP framework (Fig. 1/Appendix 2) served as a scaffold to call out a set of relationships to construct a streamlined model rooted in our specific data. Appendix 2 displays each step of the process, with coding elements and relevant excerpts. Analytically, the need to construct this streamlined model to capture important findings emerged organically after indexing the data with the elements of the TALiP framework. SF had primary responsibility for the framework analysis, with authors VM and JAL providing context to ensure the activities of the biochemistry CoP were correctly reflected. As SF engaged in analysis and co-authors VM, JAL, and JEL provided feedback on the analytic process of summarizing indexed data, we found that mapping the codes to represent specific areas of the TALiP framework muddied the overall message of the data. Therefore, we chose to use this analysis step to emphasize the process embedded in the TALiP framework rather than the individual elements. The resulting visualization abstracts relations among key elements of the TALiP framework into a system of pulleys (Fig. 1), representing the dynamism inherent to assessment literacy in practice. Pulleys serve as a metaphor to capture the interdependence of the selected elements of the model. In the following paragraphs, we will use specific areas of the model, identified as "landmarks" to describe the process of PPL's development as an assessor in the specific case of her assessment of PBI in her Spring 2016 biochemistry course. The analysis and findings were also shared with PPL for member-checking and any additional insights into the data that she could provide. Additionally, SF chose an asset-based lens to keep the focus of the study on understanding what PPL accomplished rather than on what she failed to do.

Findings

Prior assessment literacy in practice is relevant to PPL's current practice. It is important to understand what different factors influenced PPL as an instructor and an assessor because those things are linked to one another as seen in the model (Fig. 1) where the knowledge base is essential for building assessment literacy. Prior to her membership in the biochemistry CoP, PPL received her disciplinary training and doctorate in biochemistry. She was a highly experienced instructor with over 14 years of teaching experience, the bulk of which was in undergraduate biology. A hallmark of PPL's teaching practice is

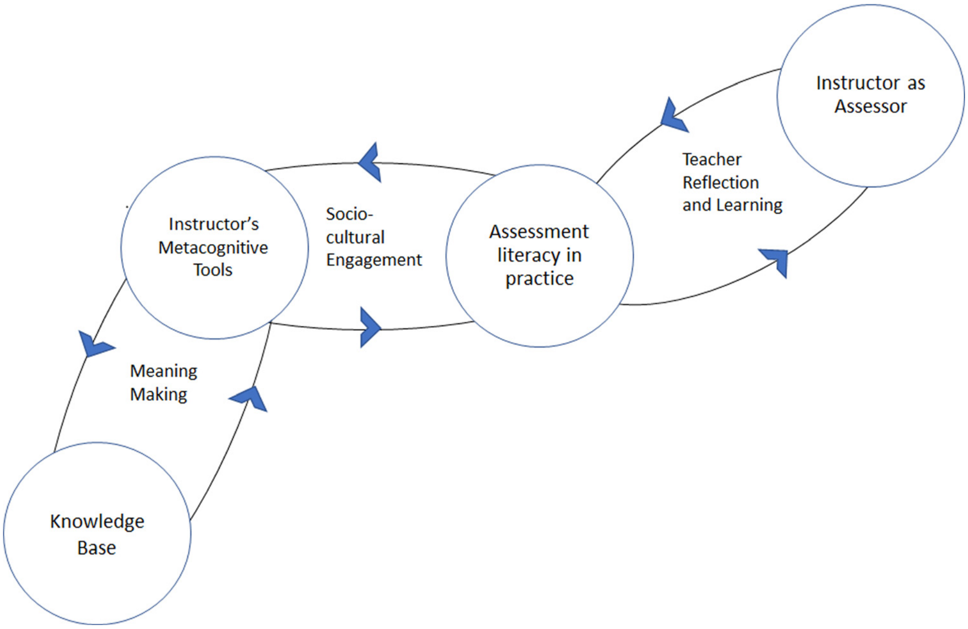


Fig. 1 Model derived from the framework analysis performed on the data of this study using the TALiP framework as a base. This model is used in the results to represent the dynamism present in PPL's assessment literacy in practice.

that she continually sought interactions with knowledgeable colleagues. Interactions with her colleagues and, in particular, her mentor during her postdoctoral study, influenced the "nuts-and-bolts" of her knowledge base and pedagogy.

I think pretty much I just didn't know anything [laughs] and so there just [...] a lot that they helped me learn [...] I mean I had done presentations before and I had actually taught some as a grad student but, I think everything I knew really was about how to present material. And so, I learned tons from them about just: course design and administration and assessment and you know, how to have fair policies and yeah, how to think about rigor and what you're trying to accomplish. (Interview 3)

It was also during her time as a postdoctoral fellow that PPL was introduced by a colleague to her first CoP: The National Center for Case Study Teaching in Science. Participation in the community influenced PPL to use case studies in her teaching from the very beginning of her career as a faculty member.

When I came to [University], I decided that I was going to bring more of that case study work in the lecture component in the course. So, I started lecturing less and doing case studies more. I did have interactive lectures using clicker questions, or some other kinds of questions as means of discussion. So, for that period of four years [before biochemistry], I was lecturing some, using more cases in class, and using bad clicker questions. (Interview 3)

The reflection on the perceived quality of her formative assessments, clicker questions, is an indication of how PPL uses her conceptions about the purpose of assessment to evaluate assessment tools she uses in her teaching practice. Her awareness of her teaching practices primed her to look for places where she could enrich her instructional and assessment practices. Another example of her awareness happened when she was observed by a colleague, which allowed PPL to recognize a dissonance between her conceptions of teaching and her teaching in action.

I just knew like – "yeah this isn't what I really want to be doing." [...] I knew enough about active learning, about how to teach, [...] and what I was doing was not aligned. (Interview 3)

PPL's reflection on the dissonance between her enacted teaching practice and her conception of herself as an instructor along with her interaction with colleagues led to PPL being referred to the biochemistry CoP by a colleague. She was receptive to that invitation as an opportunity to integrate new knowledge and tools into her knowledge base. Simultaneously, PPL's research was focused on problem solving skills and behaviors of students in introductory biology and biochemistry courses, as described in the following grant proposal excerpt:

Accomplishment of [of this research grant proposal] is expected to pinpoint students' correct and incorrect ideas about protein structure and function and metabolism. In addition, it will isolate key problem-solving steps that support or detract from domain-specific problem solving. [...] These outcomes [of the grant] are important because they are prerequisite to designing effective instructional interventions and assessments (Grant Proposal, 2013).

Because of the focus on designing effective assessments, PPL's research also influenced her assessment knowledge base and her epistemological beliefs about learning and student understanding. After years of teaching and research, PPL was poised for transformative change. As we shall see in the following paragraphs, this opportunity to change her teaching practice uniquely situated her to also enrich her assessment literacy. PPL, therefore, represents a "best case" to see this development in a compressed time frame, allowing us to illuminate critical elements of development.

Engagement in community meaning making strengthens assessment literacy (Landmark 1). Landmark 1 illustrates how assessment literacy in practice builds on foundational assessment tools and views on assessment, each of which can be enhanced and integrated in the sociocultural context of a CoP (Fig. 2). In summer 2013, PPL attended her first workshops held by the biochemistry CoP which used threshold concepts as a guiding framework to change instructional and assessment practices in the undergraduate biochemistry classroom. The workshops were an extension of a previously established biochemistry education CoP, which had worked together to encourage the use of evidenced-based instructional practices (Murray *et al.*, 2011) and to design assessment instruments (Villafañe *et al.*, 2016). The sociocultural context of the CoP enabled generation of new collective knowledge and allowed members of the community to inform each other's conceptualization of disciplinary content knowledge.

There was lots of time for conversation about you know the "stuff" that you teach in biochemistry, right? And lots of conversation about what's difficult for students, how an activity can facilitate learning, how assessments can help students to show what they've learned. But you know the conversations were really pivotal to me because I didn't have that pedagogical content knowledge. I mean I had had lots of biochemistry coursework, but I had just not taught it. So, here I was talking to people who had years and years of experience, so it just gave me a real leg up going into that [teaching biochemistry].

PPL recognized that she wanted to increase her understanding of biochemistry pedagogy. Through that recognition she used her openness to gain new insight into biochemistry to deeply engage with her biochemistry CoP peers. She worked with peers in the community to identify threshold concepts for biochemistry. PPL and the other CoP members used their collective experience, disciplinary knowledge, and knowledge about the purposes of assessment to develop instructional and assessment tools related to the threshold concepts. Through collaborative consideration of assessment practices related to specific biochemical concepts, workshop participants co-created shared standards related to assessment.

After these workshops, PPL explicitly and intentionally integrated a specific threshold concept in her classroom: the physical basis of interactions (PBI) (Table 1). The concept of PBI deals with the electrostatic interactions within and between molecules that result in the structure–function relationship of

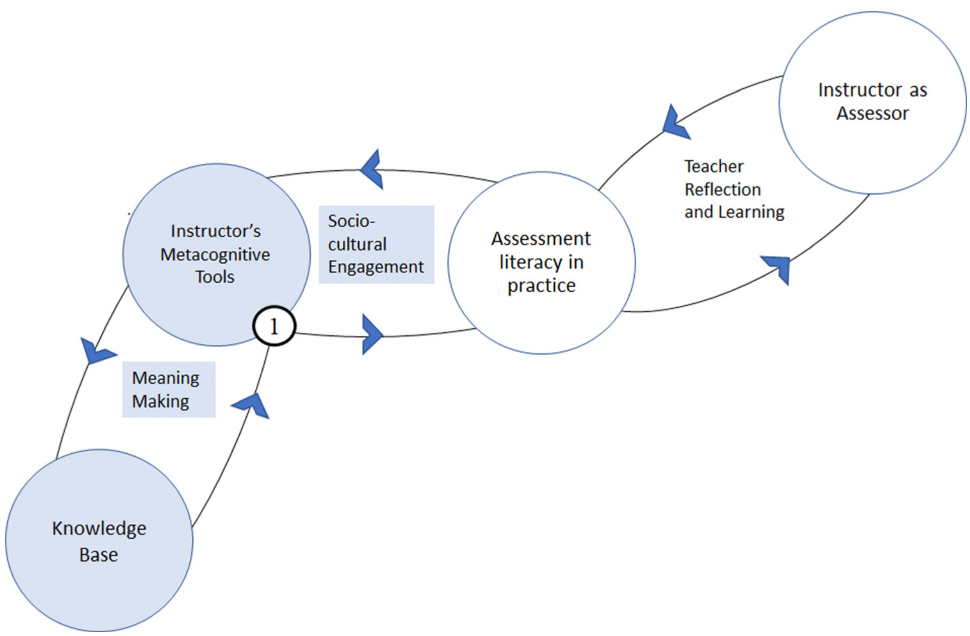


Fig. 2 Community engagement strengthens assessment literacy. Landmark 1 illustrates the ways in which PPL's conceptions about assessment were shaped by encompassing the dynamic interaction between her assessment toolbox and her participation in a CoP.

Table 1 Reproduction of the community-determined definition for PBI that was used by PPL in her classroom (Loertscher *et al.*, 2014)

Threshold concept	Knowledge statement	Biochemical ideas are unlocked once this concept is understood	Connections invisible prior to deep understanding of the concept
Physical basis of interactions	Interactions occur because of the electrostatic properties of molecules; these can have full, partial, and/or momentary charges.	Similarities between different types of interactions become clear. Although interactions are given different names, they are all based on the same electrostatic principles.	A core biochemical principle is that structure governs function. Correct understanding of noncovalent interactions is essential in integrating structure and function.

biomolecules. PBI is different from the other threshold concepts in that the concept of PBI is prerequisite knowledge for understanding structure–function relationship of proteins; without PBI knowledge it is extremely difficult to fully conceptualize most biochemical processes (Loertscher *et al.*, 2018).

PPL's connection to the community grew with every subsequent biochemistry CoP workshop she attended. Activities at the biochemistry CoP workshops included presentations on relevant DBER findings, development of new evidence-based instructional materials, and development of assessment tools to be used in instruction and research. As such, this CoP was motivated to advance assessment literacy through providing a space for biochemistry instructors to discuss their teaching experiences, techniques, knowledge about content and students and assessment practices. PPL engaged with community members who taught general chemistry, which provided a different perspective on student prerequisite knowledge of PBI as compared to her experience teaching introductory biology. PPL realized that many of her community peers viewed biochemistry from a different lens than her own:

Many of those [community members] came to biochemistry from a chemistry perspective as opposed to biology perspective. So, that really opened my eyes from the start, to that distinction because, you know the way biologists talk about biochemistry and the way chemists talk about it is quite different. [. . .] So gosh, I can remember, [chemistry colleague], we had this conversation about, well we had a couple [conversations], but the one that I remember is we were – I was discovering this whole issue that we talked about names for nonpolar interactions, right? You know I had always thought of them as van der Waals, and they're [chemistry colleague] like "we don't even use that term." And they were [also] talking about dipole–dipole, ion–dipole and I was like "what are you talking about? " I remember [. . .] she sat down and she like wrote all this out for me, how she teaches it, and I remember going back to my room and like looking up all of these things and I was like "Wait a minute, [colleague], this is not – this is not how biologists do this." So, I was explaining to her "Okay. I've taught this before in an intro bio course and we [biologists] do not use this language that you're telling me." (Interview 3)

The co-constructing of knowledge in a community is a complex, dynamic process of pulling together not only the other participant's pedagogical and assessment knowledge but the synthesis of these experiences into new community knowledge. Interactions within the CoP enabled the collective meaning making and that typically happens within a CoP (Wenger, 1999). When she encountered her chemistry colleague's difference in perspective and terminology, PPL realized that there was an alternative way to conceptualize PBI compared to her perspective from teaching introductory biology, and this influenced her disciplinary knowledge about biochemistry. PPL's receptiveness to broadening her view of how biochemistry disciplinary knowledge is informed by disciplinary training, indicated that a willingness to reflect on one's own experiences and reconcile them with peers is essential in the co-construction process. The broadening of PPL's view allowed her to "cross the threshold" herself in her understanding of a threshold concept like PBI and how this concept can be viewed in different ways. Her realization allowed for change in how she approached the topic of PBI with her students and enhanced her confidence to integrate community knowledge about threshold concepts in her classroom practice. This integration is further explored in the next landmark.

Assessment in action is influenced by socio-cultural engagement with the CoP (Landmark 2). Landmark 2 explores how PPL's 'assessment literacy in practice' (Fig. 3) was influenced by

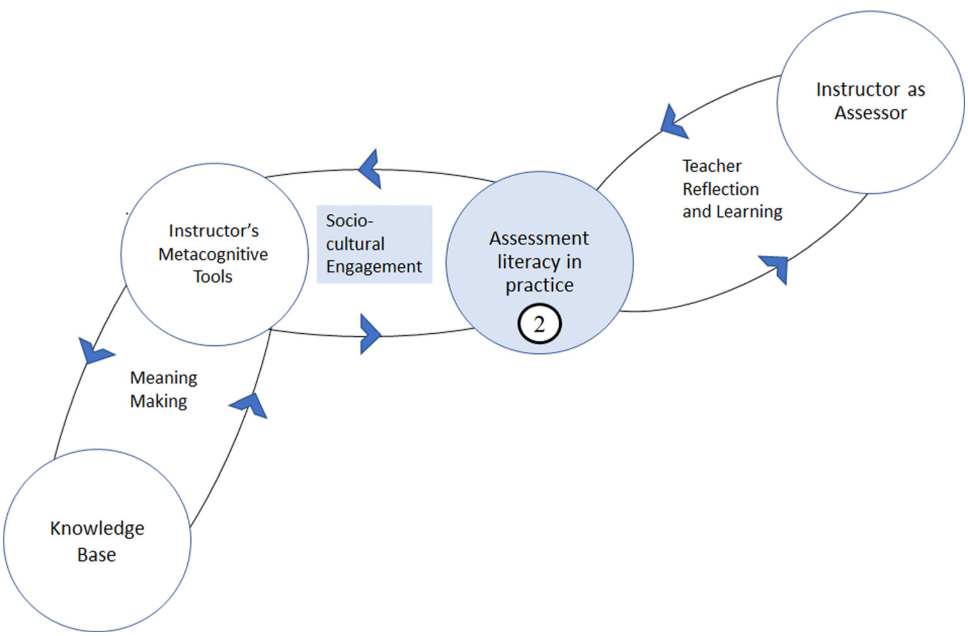


Fig. 3 Biochemistry CoP influences choices made during assessment literacy in practice in the classroom. Landmark 2 in the dynamic representation of PPL in the teacher assessment literacy in practice (TALiP) framework.

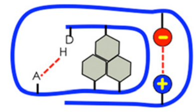
the CoPs work on the PBI threshold concept. Assessment literacy in practice is a dynamic process that involves developing and enacting assessments by drawing on assets from an instructor's knowledge base and beliefs. The assets utilized are filtered and impacted by sociocultural engagement. This landmark is defined by the ways PPL adapted the biochemistry-focused assessment knowledge she gained from the community into assessment literacy in practice. Members of the biochemistry CoP, including PPL, worked collaboratively to draft assessments related to threshold concepts, spurring the rich conversations about PBI that PPL experienced.

I think the role of noncovalent interactions and biomolecular structure [...] that's just a foundational idea of biochemistry. So, whether my students—my students are mostly not going to become biochemists, but they are going to get credit for a biochemistry course, and I do want them to come away with some foundational ideas about the discipline. (Interview 2)

The influence of the community-built knowledge of threshold concepts begins to show in instructional choices implemented by PPL in her classroom. The value that PPL places on the community's knowledge and tools can be observed when PPL speaks about integrating specific threshold concepts into her teaching practice. Another way this is evident is through changes in PPL's course learning objectives over time with respect to PBI (Fig. 4). In the initial stages of PPL's membership into the biochemistry CoP, her learning objectives related to noncovalent interactions were broad and emphasized memorization of key ideas. As PPL continued to work with the CoP for multiple years, her course objectives became more focused on specific aspects of PBI. The shift in her using terms like "weak forces" to "electrostatic" and to then explicitly using "Coulomb's Law" is another example of her integrating the chemistry lens into her conception of PBI in biochemistry. PPL's revised course objectives lay the groundwork to refine the backwards design of PBI in her course (Fig. 4).

PPL's assessment literacy was further developed through her examination of existing teaching materials to identify ways to address students' different understandings of PBI. In spring 2016, PPL modified instructional materials to build on the concept of PBI. During that semester, she adapted her existing case studies to guide her students through learning PBI concepts by first having the students examine amino acids and their electrostatic interactions and then building to the larger picture of protein structure and interactions within proteins (Fig. 5; Case study available upon request from author PPL).

Even though hydrophobic amino acids tend to be buried in the interior of a globular protein, polar neutral and polar charged amino acids locate in the interior of a globular protein as well. Look back at figure 4.26 to verify this for myoglobin. Let's look at a model of a generic globular protein (shown below) to see how this can happen.



A. Which amino acids are on the interior of this protein because of the hydrophobic effect? Label them and state what type of non-covalent interaction is occurring among them.

B. Based on this model, what allows polar neutral and polar charged amino acids to locate in the interior of a globular protein? Name the types of non-covalent interactions that are occurring between these amino acids. Reference section 2.3 of your textbook.

Fig. 5 Example question from a case study in PPL's class that focuses on PBI concepts related to amino acid interactions in myoglobin (Image credit: Zac Wood at University of Georgia; Case study available upon request from author PPL).

The integration of PBI into enacted teaching and assessment practices illustrates the importance of the biochemistry CoP for PPL and her further integration into the community. This integration allowed her to begin to make sense of PBI within her own experience as an instructor and assessor. Her established habit of interaction with the DBER literature also deepened her professional content knowledge base with respect to PBI and spurred her to draw on her knowledge of assessment design for the course exams.

The [biochemistry CoP's] work made me aware of the target and concurrently I was understanding that data for myself and also reading literature from other people. So, the [biochemistry CoP] literature but related literature that shows these difficulties. So, I just think that all of that together, kind of tuned me into what to expect from students. So, I knew there were problems there. So, I was just looking for ways to really pull out the difficulties from students, to give them questions that would force them to reckon with that. (Interview 1)

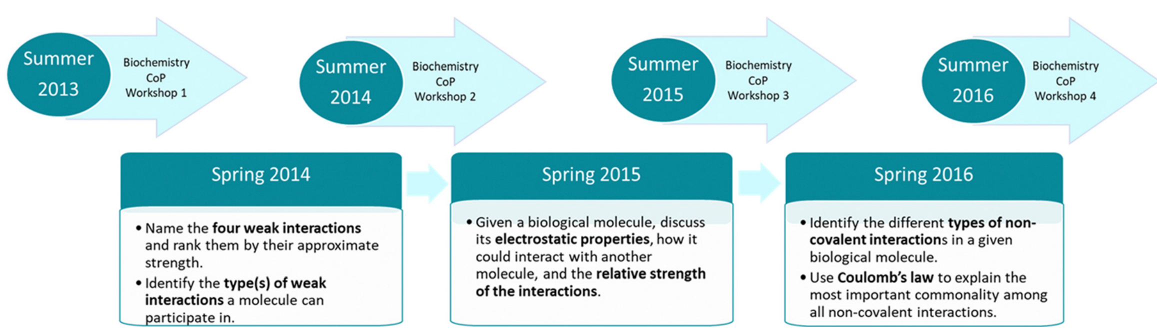


Fig. 4 The learning objectives from the first unit in PPL's curriculum that align with the concept of PBI changed over time from Spring 2014 to 2016. Integration of threshold concept specific language is bolded for emphasis.

Although PPL's course exams emphasized topics covered in each unit, her exams were also cumulative with regard to some concepts. Based on the CoP's work on threshold concepts, PPL stated that she ''*had targeted PBI as something that I wanted to carry through all of the exams*.'' PPL's choice to target PBI as a cumulative concept on exams demonstrates a fundamental precept of assessment literacy: alignment of assessment practice with instructional choices. The choice to focus on one TC rather than all five of them is an indication that along with aligning her practice with her instructional choices this involves compromise. PPL decided that PBI was the most important to focus on and to advocate for its continued assessment with her co-instructor in the course. This process of integration and making of strategic choices and compromise is essential to assessment literacy in practice.

Teacher reflection and learning can happen iteratively within one term (Landmark 3). In Landmark 3, PPL engaged in planning exam questions based on desired PBI learning outcomes, administered an exam to students, reflected on student difficulties, and used that information to inform questions on subsequent exams with the goal of better eliciting student thinking about PBI (Fig. 6). Refinement of her assessment practice was influenced by PPL's reflective nature and her focus on desired student learning outcomes related to PBI. PPL viewed instruction as a way to ''support the ability [of students] to stretch on the exams.'' Through her work with the community and her assessment knowledge, PPL came to view structure–function relationships in proteins and other macromolecules as an excellent context for the development of rich questions. For example, her second exam contained a question that asked students to identify a molecule in the cytoplasm that would interact with glutamic acid in a protein shown in an image in the question.

As shown in Fig. 7, the question includes a figure that relates to the content in the course and is integrated into the introductory prompt of the question. While the question has an open response format, it is structured in a stem-and-leaf manner. Part A prompts students to think about the indicated amino acid and to identify a molecule in the cytoplasm with which it could interact. In part B students are then prompted to demonstrate their understanding of PBI by creating a visual representation of the interaction between the amino acid and the molecule they identified.

I think this is what happened: so, my colleague and I—my co-instructor who taught the latter part of the course, we had decided that, you know, everything would be cumulative. That what we would do is that the cumulative portions would come in the short answer, and I had targeted PBI as something that I wanted to carry through all of the exams. And so, I created the exam two question I think um, it was not really hinging on exam 1. Rather it was just meant to be another iteration of this, you know, can you draw these noncovalent interactions? So, I think what happened is: their [the students] very poor performance on this item, so it wasn't exam 1, but it was their poor performance exam 2 that prompted me to keep going with [the concept through] exam 4.

When the responses to this question became part of PPL's assessment literacy in practice as part of her second exam, she faced an unexpected result. Fifty-three percent of the students answered part A with PPL's anticipated response: water. The other 47% of students gave a range of answers, including free or bound amino acids, glucose, and other molecules.

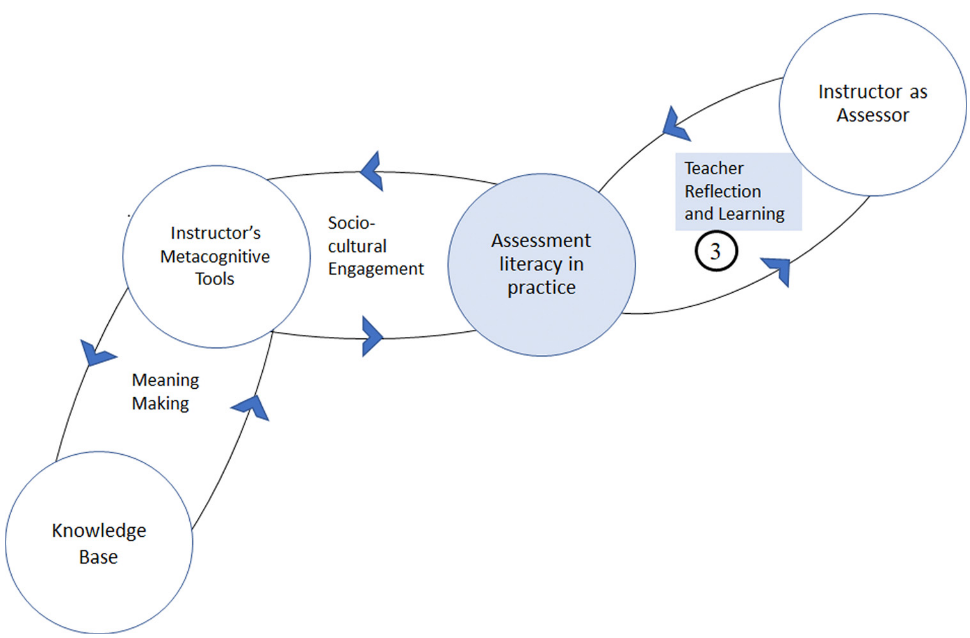


Fig. 6 Reflecting on assessment outcomes can lead to refinement of assessment informed by lessons learned in practice. In Landmark 3, PPL engages in the teacher reflection and learning cycle to further align her assessment literacy in practice related to PBI objectives.

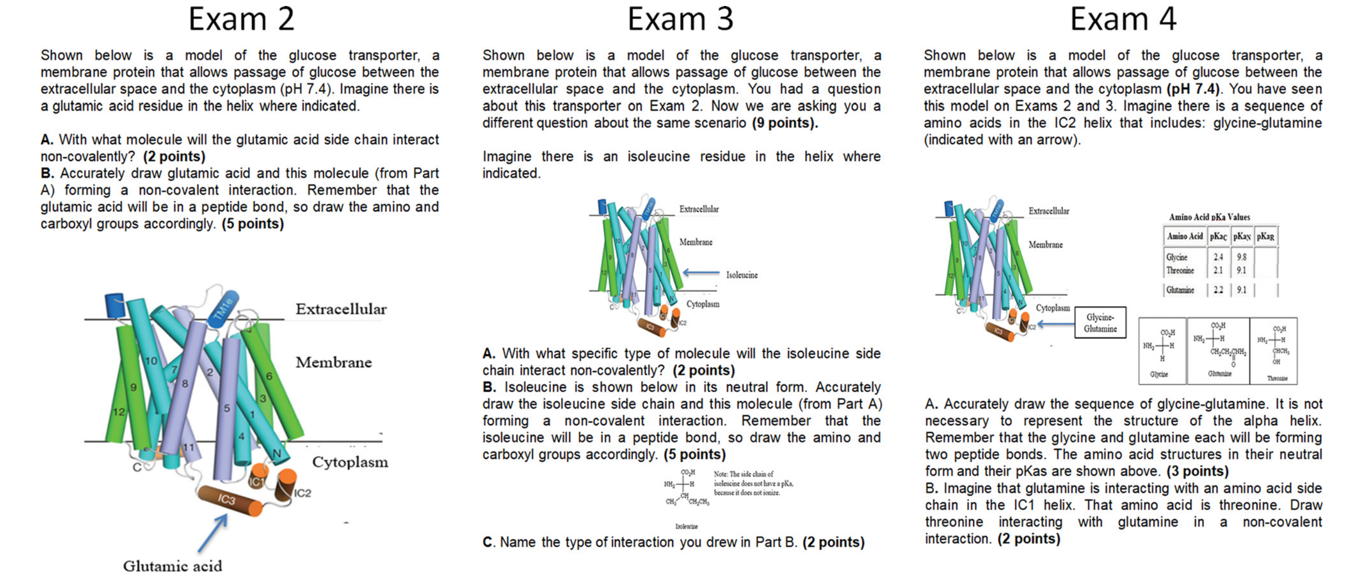


Fig. 7 Questions from exams 2–4 that PPL gave to her students to assess PBI (protein image reprinted with permission, Deng *et al.*, 2014).

Okay so with 'Part A,' I wanted them to recognize that the solvent was water or polar at least and I was stunned that they couldn't do that. [...] I mean what really tripped them because they really didn't know what [part] A was about; they didn't—they could not imagine what this molecule was. (Interview 1)

Although PPL recognized that the open format of the question allowed for responses other than water, the degree to which many responses were unanticipated enabled her to see that something she viewed as ubiquitous appeared to be missing or opaque to students. This unanticipated outcome was troubling to PPL, and that motivated her to reformulate her approach to assessing students' understanding of PBI while reaffirming her decision to include PBI required cumulative concept on multiple exams.

Over the course of this semester, PPL's exams showed a noticeable evolution with respect to assessing student understanding of PBI. After the surprising result from the PBI question on the second exam, PPL's reflection and learning cycle is obvious as she discusses her plans for this PBI question on the next exam. She reflected on the specific difficulty she believed the students encountered with the initial formulation of the question, and she intended to reformulate the question to bypass this difficulty. In this reformulation, PPL taps into knowledge and tools from the CoP to posit PBI as a unitary concept, regardless of whether the interactions are polar or nonpolar.

I wanted to give them the same model again and say 'look, you've been asked about this before. I'm just going to ask you a different question about it this time.' [...] now I gave them [indicates with mouse on the figure] an isoleucine residue on a helix that's embedded in the membrane. [...] I knew that this would be an easier grab for them to recognize that in the membrane we were talking about nonpolar molecules. I was pretty confident that, that was going to work and so now I was just getting them to think about—it's kind of the same type of question [as exam 2 Q22] with nonpolar interactions. (Interview 1)

With her focus on another aspect of electrostatic interactions, nonpolar interactions, PPL attempted to support students to show their understanding of the interactions between the protein and the membrane rather than with the cytoplasm or exterior environment of the cell (Fig. 7). Instead of asking students to identify a specific molecule with which an amino interacts (Fig. 7), she instead asked students to state a molecule type. She included more visual aids in the form of a line drawing of free isoleucine and a reminder that it has a non-ionizable side chain due to pK_a values. She also prompted students to think about the previous exam question and explicitly told them that they will be considering another aspect of the same protein. The fine-tuning shows that PPL was thinking about how to best prompt students to consider interaction type and was aware of possible student difficulties. When enacted in exam three, the tighter focus on the PBI was successful in prompting a majority of students to recognize that the nonpolar isoleucine would interact with the hydrophobic tails of the membrane. However, another issue arose.

So, there was a point at which I figured out that they didn't know how to draw a peptide bond [...] It just became clear that they didn't know that. (Interview 1)

Students' representations of isoleucine as a free amino acid and not part of a peptide chain indicated that they were not thinking about the indicated amino acid as being a part of a whole protein interacting with the membrane. Approximately a third of the students answered this question by showing partial or full peptide representation. The non-use of molecular representation of the peptide indicated that there was a need for an intervention to better support students to show their understanding of biomolecular structures. These representations are of particular importance to the PBI concept because of the central role of noncovalent interactions in structure–function relationships in macromolecules. This recognition meant that PPL had to again reach into her assessment literacy toolbox and draw another approach to assessing students on the fourth and final exam.

So, with exam four, the final exam, I gave them yet again the same scenario and I had told them: 'you're going to have to draw a peptide bond', and I then wanted to go back one more time to this idea of 'can you draw the non-covalent interactions?' (Interview 1)

The newly drafted question for the final exam (Fig. 7) follows the same scaffold of the other discussed exam questions. As with the change from the second to the third exam question, she tightens up the assessment by making the question more specific. The intentionally specific prompts allow students to demonstrate their ability to show appropriate representation of amino acids in a peptide, based on their struggle with the question in Fig. 7. The next part of the question targets students' understanding of PBI by prompting students to recognize that the side chains of the given amino acids can then interact with another amino acid within the same protein in a non-covalent manner. These two parts of the question are supported with a table of pK_a values for three amino acids listed in the question in addition to the visual representation of the free amino acids and the protein. The results of this third question, showed that 75% of students were now able to demonstrate full or partial representation of the indicated amino acids in a peptide bond and 80% of students showed correct electrostatic interactions between the glutamine and threonine.

PPL is a reflective assessor (Landmark 4). In Landmark 4, we consider how PPL's identity as an assessor was impacted by what she learned from taking action to iteratively implement assessment changes and reflected on the outcomes of each subsequent assessment. As we saw with other landmarks, all the elements of the framework work together to play a larger role in how PPL views herself as an assessor (Fig. 8). This reinforcement can be seen in her existing depth of epistemological beliefs about the nature of assessment and learning is evident in the way she discusses her original intent:

I think I was really imagining that if I could get them to draw out these things that, that is a deeper level of knowledge than just like citing the definition or even—you know I have some multiple choice that will say 'what will this side chain interact with' and 'what can this one interact with' but I felt like the drawing and sometimes explanations will provide a bit more evidence. [...] I really look at assessment as informing the student as well, so kind of forcing the way that they prepare. So, I was thinking that it would be information for me but also information for them like, 'these are things I need to understand.' (Interview 1)

PPL uses her exams as summative assessments to provide a grade to her students and as formative assessments to help her understand students' difficulties and provide feedback to students. Her care and concern for students is evident. PPL posted a detailed rubric for each assessment in the course management system. Specifically for exams, she reviewed the detailed rubrics during class time. PPL further extended the opportunity for students to engage with her in one-on-one meetings for feedback. This indicates that she saw herself not only as someone who drafts assessments but as an assessor who thinks about all aspects of feedback from assessments.

I mean I think what I've recognized, they really just need [...] some work thinking about [...] these interactions in the context of biological molecules. 'So, let's just forget about water and let's just think about, you know, the interactions themselves. I mean I know you have to assume a solvent but, but let's assume you know a neutral pH and then just start there.' I mean they need that work and then like when you muddy it with the water and kind of give them another layer of things to think about then, I think it's just hard for them to know 'Wait. Am I supposed to be thinking about the interactions or the hydrophobic effect?' (Interview 1)

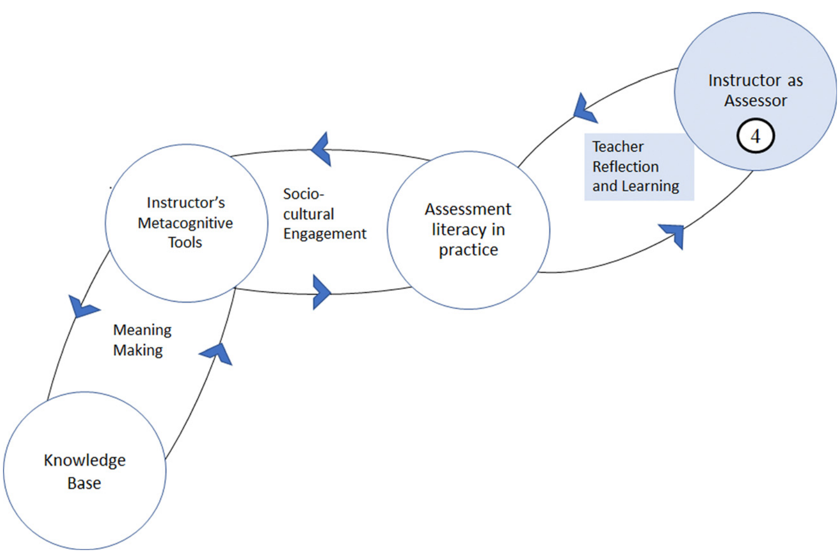


Fig. 8 In Landmark 4, PPL's identity as an assessor is the sum of knowledge, views, and assessment practices.

The reflective nature of PPL's approach to teaching and assessing is a prominent feature we observe as she refines exam questions within one term of a course. We observe her flexibility in moving from one domain of assessment literacy to another and her ability to keep working on her goal of aligning assessment of PBI. This commitment allowed her to gain more insight into not only student understandings, but also understanding of her own assessor identity. As she looked ahead, PPL thought about what could have impacted student performance on these questions and what she may need to do in her teaching to prepare students for the types of assessments she wants to use in the future.

I'm going to do this totally differently [. . .] you know sometimes you are like: 'Really? Have I learned that much since then?' [. . .] I just think this [the exam questions] is too much of a stretch from what I asked them to do in the cases. [. . .] I'll have to think about—because there's just so much—and I like the case questions, but I wonder if I can just start with these exam questions and work back from there? So, a new approach that I'm planning for my teaching is to have a central problem, or set of problems, that I want students to be able to solve every week. (Interview 1)

The process of reflecting on her assessment practices further prompted her to consider how to change instruction using case studies. As PBI was new to PPL's assessment literacy in practice it was necessary for her to examine and reflect on her assessment outcomes in order for her to identify the elements needed to integrate instruction and assessment of PBI.

In looking at these exam questions, the insight that I have gained during this interview is actually what I like them to be able to do: [laughs] is to do these questions! Right? And so that's what I'd like to build my case around and then kind of work backwards from there. Like what—what are the pieces that you really need in order to do this question? (Interview 1)

Ultimately, this process of reflection and iterative realignment is what allowed PPL to move through the different aspects of assessment literacy in practice. Reflection and connection to future teaching is something that is a hallmark of a teacher-as-assessor identity and is clearly apparent in PPL's practice.

Discussion and conclusion

How was the instructor's assessment literacy in practice influenced by membership in a biochemistry community of practice?

This work adds insight to the call by Schafer and Yezierski (2020 and 2021b) for how instructors translate knowledge of assessment into assessment practice, by showing that there are specific aspects of instructor support that helped PPL in contextualizing how to interpret her assessment data. In particular, the biochemistry CoP provided a touchstone for PPL, focusing her efforts to align her biochemistry instruction to community developed PBI objectives. In this study we observed PPL grow as an assessor over the course of a semester as she drew upon her assessment knowledge and the collective assessment literacy of a community of practice. This observation supports the suggestions by Looney *et al.* (2018) and DeLuca *et al.* (2021), that instructors' beliefs and their cultural context are important elements of teacher assessor identity. PPL engaged in the biochemistry CoP, which was specifically developed to build pedagogical support for threshold concepts, including PBI. The biochemistry CoP provided space for community members to contribute their PBI knowledge to the research literature, reinforcing the status of PBI as an important concept for biochemistry instruction (Murray *et al.*, 2011; Villafañe *et al.*, 2011a, 2011b, 2016; Loertscher *et al.*, 2016, 2018; Xu *et al.*, 2017, Villafañe *et al.*, 2021). PPL took her new insight about the importance of PBI back to her home institution. She redesigned learning objectives over multiple semesters to better communicate her expectations to students and to align with what she believes PBI proficiency looks like in biochemistry. PPL's intentional focus and alignment is a shining testament that instructors, with community support, can reconcile the misalignment that Schafer *et al.* (2021) characterized between learning objectives and assessment items.

Through the knowledge and perspective provided by the biochemistry CoP, PPL was able to identify a misalignment in her assessment outcome and the awareness she had of her students' content knowledge. PPL's intentional integration of PBI and her awareness of her own classroom (both pedagogy and knowledge of students) allowed her the insight to alter her assessments in a way that provided students the opportunity to show their knowledge through assessment within the same semester. PPL's development and journey as a biochemistry assessor makes her an important case to examine because not only did her assessment practices change but those changes were due to her intentional focus on PBI, a crucial concept in her discipline. Her intentional thread of PBI focused her attention to student outcomes on PBI assessments and she leveraged her reflexivity to refine an assessment item within the same semester rather than doing the refinement for the next semester. PPL then further extended her reflection on her assessments to further consider what specific PBI assessment outcomes were important to her as an assessor. She recognized that to gain the desired outcomes she would need to further align her assessments and instructional practice. This act of aligning her assessor identity with her teaching practice is consistent with a refinement in her assessment literacy.

What assets did the instructor leverage to assess student understanding in a biochemistry context?

Understanding instructor assets related to assessment literacy is an important gateway into providing better professional development for faculty and ultimately positive student learning outcomes. PPL leveraged multiple assets as she continued to develop her identity as an assessor during this study. These assets include knowledge and skills, such as content knowledge about biochemistry and her DBER-influenced awareness of the structure and purpose of assessments, as well as beliefs, such as her commitment to assessments as opportunities for students to learn. Timely community support influenced her determination to focus on PBI in her biochemistry course and her perceptions of student assessment outcomes, but overall, what is most striking is her reflexivity in regards to her teaching and assessments and her willingness to change. As Offerdahl and Tomanek (2011) have noted, while instructors may change their assessment practices that does not necessarily mean an immediate change in their teaching practice or, indeed, any change at all in their teaching practice. What is noteworthy about PPL is that she monitored students' performance with respect to PBI and responded in that same semester by altering both instruction and assessment practices. While studies have shown that instructors often intend to align their objectives with their assessment, actual alignment is a skill that needs to be developed (Schafer *et al.*, 2021). PPL's action of modifying assessments within a single semester shows one way that this skill development can occur. PPL applied her metacognitive and analytical skills to integrate the information that she gathered from her assessments and applied her insights in two directions: one as a method of informing students' understanding of biochemistry content and two as a way to use the insights to critically look at her assessments themselves. Such metacognitive awareness has been indicated as an area of further examination (Wang, 2020; Herridge *et al.*, 2021).

In what ways did the instructor's learning and reflection on assessment outcomes influence her identity as an assessor?

PPL demonstrated the dynamic nature of teacher assessment literacy in practice, as seen in the TALiP-derived pulley model, as she iteratively reflected on her assessment and teaching practices and enacted changes to both. She used her analytical skills to integrate the information that she gathered from her assessments and apply her insights into her practice through intentional focus on PBI. PPL's awareness of her own classroom (both pedagogy and knowledge of students) enabled her to alter her assessments within the same semester, ultimately providing students with ample opportunity to display their PBI knowledge. As PPL examined her assessments, she further reflected on how she as the assessor can align her instructional practices to better support student understanding and outcomes on assessments. Her assessor identity was strengthened through this process, as she honed her vision of exams as opportunities to gain understanding of students and to express her own values regarding the importance of PBI for biochemistry.

Implications

Summarized below are implications from this study for instructors, professional developers, and researchers. These implications work toward expanding understanding of assessment literacy and instructor assets.

Instructors should embrace experimentation and reflection

Assessment is commonly discussed as a way for students to develop but not so much as a way for instructors to grow as assessors. We see with PPL that treating all assessments as formative to yourself as an instructor can allow for flexibility and growth. With this view, an instructor has permission to reflect on enacted assessments, how they function, and how they may or may not be meeting intended course goals. Instructor assets develop over time, and that development can be expressed through assessment practice. Refining learning objectives and revising assessments to continue to develop greater awareness of student thinking makes sense. Although it can be daunting and sometimes discouraging to critically engage in a process to improve assessment practice, working with a community of colleagues can provide support and direction for growth as an assessor, with the ultimate goal of better supporting student learning.

Professional developers should aim to leverage assets

The TALiP framework is beneficial for building greater understanding of enacted assessment practice, but it can also be leveraged to scaffold professional development for faculty. In this study, the TALiP framework helped elucidate that assessment literacy and enacted practice can be enriched even for an experienced assessor like PPL, who needed several iterations of assessments in order to uncover student understanding. This important feature of assessment practice – iteration – needs to be supported and encouraged in professional development. Normalizing growth in this way is an important aspect of asset-based professional development. Positive social reinforcement is more likely to encourage change (Bouwma-Gearhart *et al.*, 2014; Gast *et al.*, 2017; Gehrke and Kezar, 2017). An organized CoP, such as the one PPL participated in, can be an important facilitator of growth through the valuing of process (Gerhke and Kezar, 2017). A CoP can provide asset-based accountability, not focused on getting things right, but rather on a commitment to exploration, growth, and sharing.

Researchers should take an asset-based view of instructor development

The TALiP framework of Xu and Brown (2016) is helpful for understanding assessment literacy and assessor identity, and how both can be observed in practice. While our study of PPL's development as an assessor drew specific aspects of this framework together into a streamlined model, the larger framework provides a wealth of potential foci for continued research. Gaining more understanding of the range of assessment approaches and assessor types, the implications of cultural and institutional contexts, and the multiplicity of assessor goals can only help to push the field forward. Regardless, it will remain beneficial to view instructors through a lens of strengths rather than through their knowledge "gaps". By focusing on describing what "is" rather than on comparisons to an abstracted ideal, researchers will elucidate more nuanced understandings of assessment literacy and practices (Potari, 2021; Superfine 2021). Investigating assessor identity and acknowledging instructors as active agents in their own identity development can alleviate some of the natural tension between what "is" and what "could be." A prime example of this tension can be seen in a paper published by

Harshman and Yeizirski (2016) in which they acknowledge, "*[w]hile our study may seem to paint chemistry teachers' ability to design and interpret assessments in a negative light, we do not believe that these teachers are at all 'unable' to do this.*" Likewise, the authors of this paper have not been immune to the deficit frame, but we strove to focus on describing the assets that PPL leveraged to work towards her self-identified assessment goals. The TALiP framework has been very helpful to support this focus.

Limitations

As with all naturalistic and interpretivist approaches, it is important to consider the context of the data when considering the potential for transferability. This data and analysis focused on a single individual, and results will likely differ for other individuals and in other contexts. This work, however, is an important step for gaining a deeper understanding of assessment literacy in practice through examining a "best-case" scenario. It also provides insight into different types of assessment assets that faculty may bring to bear in their classrooms, with a specific focus on improving assessment practice for a threshold concept in biochemistry. Another aspect to note is that this study focused on a CoP external to the institution and not those within the institution. We acknowledge that more than one CoP can play a part in an instructor's assessment practices and we look forward to continued research into this area. We also recognize that this study focused on the instructor and her interpretations of student responses to assessment items. The instructor perspective is one facet of assessment in practice. The other important facet is the student perspective. We believe that future studies would benefit greatly from gathering student voices.

Conflicts of interest

There are no conflicts to declare.

Appendix 1. Interview protocols

Interview 1 questions: interview about assessments and TCs

1. How do you define PBI? How did you utilize the PBI construct map? Do you have a copy of the map that you used?

2. When you were designing your biochemistry curriculum did you build it from the 'bottom-up' with the PBI threshold concept as the base or did you modify an existing curriculum?

3. Did the PBI threshold concept influence your expectations for student learning? Were your expectations for students changed by learning about the PBI threshold concept?

4. What sort of progress would you expect of students as they moved through the coursework in terms of understanding of PBI?

5. Consider Q26 on Exam 1:

1. What kind of answer did you intend to elicit from students?

2. Did that occur?

3. Was there something that stuck out to you in student responses?

4. Is there anything you would do different with this question?

5. Anything you would keep the same?

6. Did Exam 1 Q26 have any influence on Exam 2 Q22 when you were writing it? If so, how did it factor in your writing for this question?

7. Consider Q22 on Exam 2:

1. What were you trying to probe with Q22?

2. Was there anything that surprised you or stuck out to you about student answers on the question?

3. Is there anything you would do different with this question?

4. Anything you would keep the same?

8. Consider Q26 on the final exam:

1. What were you trying to probe with Q26 on the final exam?

2. It seems like it was related to E1Q26 and E2Q22, is that true? Did those questions have any influence when you were writing Q26? How so?

3. Is there anything you would do different with this question?

4. Anything you would keep the same?

9. Is there anything that you would do differently when trying to probe student understanding of noncovalent interactions between amino acids?

10. Is there anything that you would differently in regard to assessment of student understanding of PBI?

Interview 2 questions: instructor interview on teaching and TCs

In relation to the myoglobin case study/classroom activity:

1. Identify one 'big idea' (or topic) that you teach.

2. What is most important for students to know about this idea?

3. Why is it important for students to know this?

4. What do you know about your students' thinking that influences your teaching of this idea?

5. How did you put the 'big idea' into the activity? Where is it in the activity?

6. What teaching strategies did you use?

7. Why did you choose these teaching strategies to engage students with this idea?

8. What was the source of these teaching strategies?

9. How did you ascertain students' understanding or confusion around this idea after teaching them?

10. Did your implementation in the activity work as you expected?

11. What do you do if students are not grasping the material during the activity? What do you do if students did not perform as expected on the activity/case study?

In relation to the Bartter syndrome case study:

1. Identify one 'big idea' (or topic) that you teach.

2. What is most important for students to know about this idea?

3. Why is it important for students to know this?

4. What do you know about your students' thinking that influences your teaching of this idea?

5. How did you put the 'big idea' into the activity? Where is it in the activity?

6. What teaching strategies did you use?

7. Why did you choose these teaching strategies to engage students with this idea?

8. What was the source of these teaching strategies?

9. How did you ascertain students' understanding or confusion around this idea after teaching them?

10. Did your implementation in the activity work as you expected?

11. What do you do if students are not grasping the material during the activity? What do you do if students did not perform as expected on the activity/case study?

Additional questions:

1. Why do you continue to use case studies?

2. What is it about case studies that works well with your teaching style?

3. What kind of strategies do you employ to circumvent any major pitfalls for students?

4. Why do you choose to use a flipped classroom?

5. How do you structure your flipped classroom?

6. Are there other types of classroom teaching strategies that interest you? Would you think about using them in your classroom? Why or why not?

Interview 3 questions: instructor interview on community

1. When you first started teaching, what helped you the most as you established yourself as an instructor? Was there something that was pivotal in helping you?

2. Who do you talk with at your home institution regarding teaching? Why do you talk with them?

3. What sorts of conversations to you have with your colleagues about teaching? About biochemistry as a subject?

4. Can you think of a conversation that you have had with a colleague that has changed your perspective on an aspect of instruction? On an aspect of the field of biochemistry?

5. What do you think helped you the most when you first began to teach biochemistry?

6. What motivated you to attend your first core collaborators workshop (CCW)? What was the developmental stage of your biochemistry teaching at the time?

7. What motivated you to return in subsequent years?

8. What activities at the CCWs most influenced your teaching?

9. Did the ways in which the CCWs influenced your teaching change as you gained more teaching experience?

10. How did relationships you developed at CCWs affect your teaching?

11. How did participating in the identification of the threshold concepts and developing the construct map related to PBI influenced teaching your perspective on teaching? Did it affect how you implemented things in your classes?

12. How do you qualify your perceived movement from a new member to the group to a core member related to teaching?

13. Does this sense of community affect you as an individual and instructor?

14. What benefits do you think you have gained during the process of working on the current manuscript?

Appendix 2. Framework analysis process

Step 1. Data familiarization

The following table details the process of data familiarization, wherein the data was explored using initial areas of inquiry (column 1) and then the sources roughly sorted by emergent themes related to those areas.

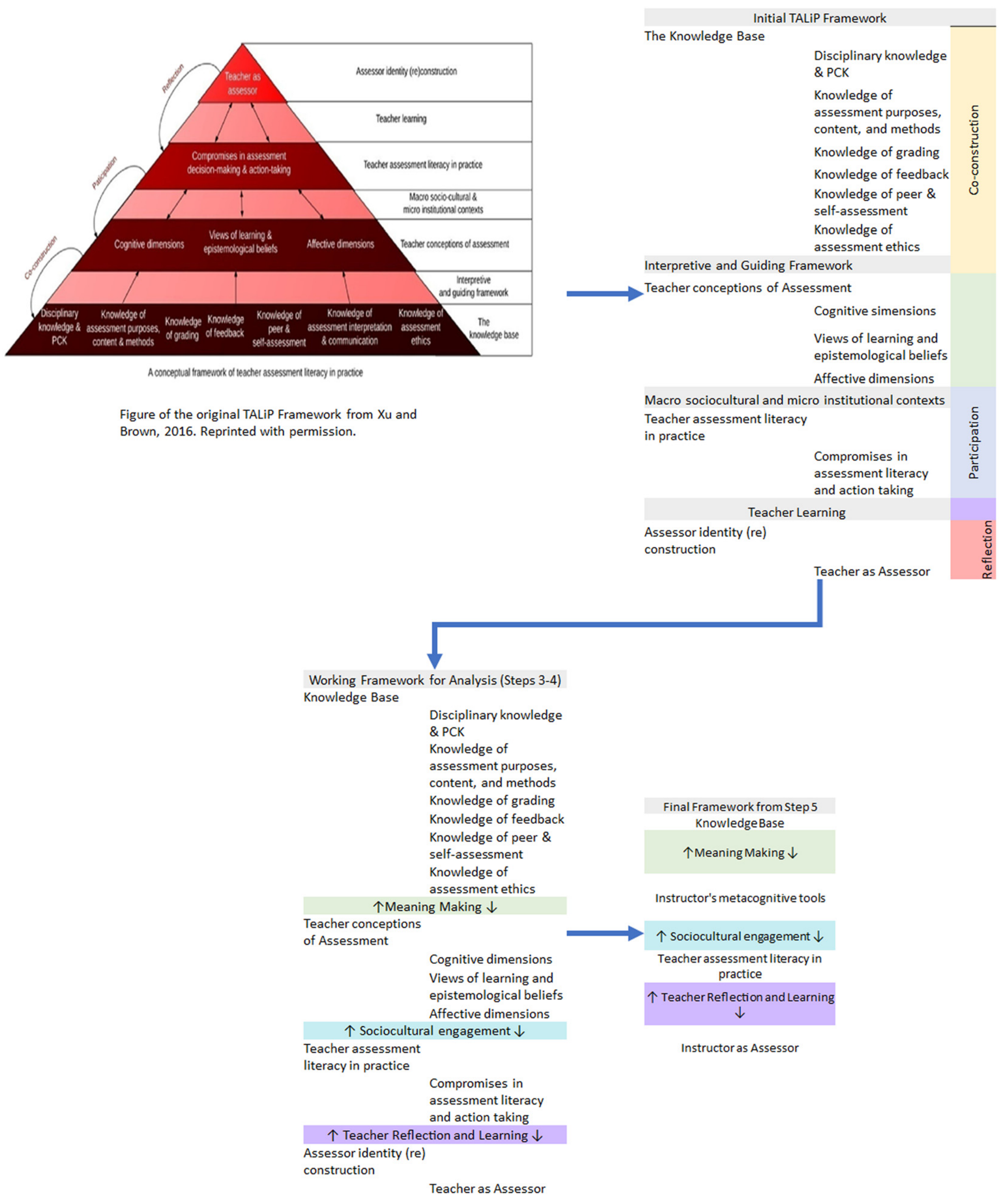
		Presence of emerging theme in data source					
Areas of inquiry	Emerging themes	Instructor interview on teaching and TCs	Instructor interview on community	Interview about assessments and TCs	Assessment items	Classroom artifacts	Workshop and other artifacts
Threshold concepts in biochemistry	PBI is foundational	✘	✘		✘		✘
	Scaffolding and backwards design easier with TC as anchor			✘	✘	✘	✘
	Knowledge of students drives content framing	✘	✘	✘	✘	✘	
Instructor perspective on assessment	Exams can be viewed as being formative and summative			✘			
	Assessment development is iterative	✘		✘	✘		
Instructor perspective on teaching	Reflexivity drives changes over time	✘	✘	✘		✘	
Community of practice	Community support influences confidence		✘	✘			✘
	Communication with community expands conceptions		✘	✘			
	Community influences teaching and assessment practice	✘	✘	✘	✘	✘	✘

Step 2. Framework identification

During the process of framework identification, it became clear that the full TALiP framework drew focus away from important patterns in the data. As the analysis unfolded, we noted the process of assessment literacy in action as important to capture in order to highlight the dynamism we were seeing in the data. This streamlining of the TALiP framework manifested as condensing certain elements of the framework so that interactions were not obscured. An example of elements that were merged is the forward and back filters/arrows between each of the level of the full TALiP framework. The ''co-construction'' arrow and the filter level ''interpretive and guiding frameworks'' were merged into a single category of ''meaning-making'' that supports a broader theme regarding how co-construction and interpretive and guiding frameworks inform each other in a recursive process.

Step 3–4. Indexing (coding) and charting the data

The following table shows the complete codebook of deductive and inductive codes used in the indexing/coding of the data. The inductive codes were generated during the data familiarization process in Step 1. The inductive codes were then



	Framework Deductive Codes	Inductive Codes Connected to Framework
Knowledge Base	Disciplinary knowledge & PCK	Learning about teaching techniques, threshold concepts (TC), Knowledge of disciplinary education research
	Knowledge of assessment purposes, content, and methods	Value of PBI, type of assessment
	Knowledge of grading	
	Knowledge of feedback	
	Knowledge of peer & self-assessment	
	Knowledge of assessment ethics	
↑Meaning Making ↓		Applying beliefs to knowledge, Valuing community knowledge
Teacher conceptions of assessment	Cognitive dimensions	Awareness/Knowledge of students, Awareness of gaps in practice, Metacognitive flexibility
	Views of learning and epistemological beliefs	Beginning of career, Knowledge/Awareness of students
	Affective dimensions	Feeling supported/ confidence, Benefits of community
↑ Sociocultural engagement ↓		Interactions with colleagues, Participation in community
Teacher assessment literacy in practice	Compromises in assessment literacy and action taking	Problem solving/troubleshooting, Use of TC PBI, using knowledge of assessment, Using community knowledge, Community influence on assessment practice
		Barriers to implementation
↑ Teacher Reflection and Learning ↓		Opportunity for change, Changes in assessment practice, Responding to student outcomes
Assessor identity (re)construction	Teacher as Assessor	Refinement of knowledge going forward, Reflection on potential changes to practice, Insights into practice

associated with elements of the TALiP framework that were used as deductive codes during the indexing (coding) process.

The following table shows an example indexing (coding) and charting process. This process starts with indexing the data and connecting the data to the framework, leading to the themes revealed from the data. A data excerpt would then be connected to the mapped landmarks in the derived model from Step 5. For brevity, this table is a truncated version of the full coded/indexed data. Analytic memo-ing was also part of the analysis process but is not included here.

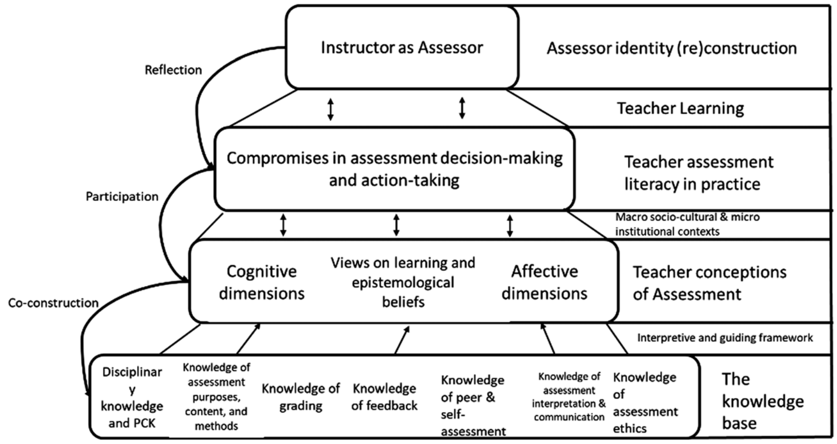
Data excerpt	Codes applied	Framework connection	Final theme	Associated landmark
I just knew like – "yeah this isn't what I really want to be doing." [...] I knew enough about active learning, about how to teach, [...] and what I was doing was not aligned. (Interview 3)	1. Reflection on teaching practice	Teacher reflection and learning	Prior assessment literacy in practice is relevant	0
Many of those [community members] came to biochemistry from a chemistry perspective as opposed to biology perspective. So, that really opened my eyes from the start, to that distinction because, you know the way biologists talk about biochemistry and the way chemists talk about it is quite different. [...] So gosh, I can remember, [chemistry colleague], we had this conversation about, well we had a couple [conversations], but the one that I remember is we were – I was discovering this whole issue that we talked about names for nonpolar interactions, right? You know I had always thought of them as van der Waals, and they're [chemistry colleague] like "we don't even use that term." And they were [also] talking about dipole–dipole, ion–dipole and I was like "what are you talking about?" I remember [...] she sat down and she like wrote all this out for me, how she teaches it, and I remember going back to my room and like looking up all of these things and I was like "Wait a minute, [colleague], this is not – this is not how biologists do this." So, I was explaining to her "Okay. I've taught this before in an intro bio course and we [biologists] do not use this language that you're telling me." (Interview 3)	1. Disciplinary knowledge and PCK 2. Metacognitive flexibility 3. Interactions with colleagues	Meaning making; instructor's metacognitive tools	Engagement in community meaning making strengthens assessment literacy	1
I think the role of noncovalent interactions and biomolecular structure [...] that's just a foundational idea of biochemistry. So, whether my students—my students are mostly not going to become biochemists, but they are going to get credit for a biochemistry course, and I do want them to come away with some foundational ideas about the discipline. (Interview 2)	1. Value of PBI 2. Views of learning and epistemological beliefs	Sociocultural engagement	Assessment in action is influenced by socio-cultural engagement with the CoP	2
The [biochemistry CoP's] work made me aware of the target and concurrently I was understanding that data for myself and also reading literature from other people. So, the [biochemistry CoP] literature but, related literature that shows these difficulties. So, I just think that all of that together, kind of tuned me into what to expect from students. So, I knew there were problems there. So, I was just looking for ways to really pull out the difficulties from students, to give them questions that would force them to reckon with that. (Interview 1)	1. Cognitive dimensions 2. Using community knowledge 3. Community influence on assessment practice	Assessment literacy in practice	Assessment in action is influenced by socio-cultural engagement with the CoP	2
I think this is what happened: so, my colleague and I—my co-instructor who taught the latter part of the course, we had decided that, you know, everything would be cumulative. That what we would do is that the cumulative portions would come in the short answer, and I had targeted PBI as something that I wanted to carry through all of the exams. And so, I created the exam two question I think um, it was not really hinging on exam 1. Rather it was just meant to be another iteration of this, you know, can you draw these noncovalent interactions? So, I think what happened is: their [the students] very poor performance on this item, so it wasn't exam 1, but it was their poor performance exam 2 that prompted me to keep going with [the concept through] exam 4. (Interview 1)	1. Compromises in assessment literacy and action taking 2. Use of TC PBI 3. Responding to student outcomes 4. Reflection on assessment practice	Assessment literacy in practice; teacher reflection	The teacher reflection and learning can happen iteratively within one term	3
In looking at these exam questions, the insight that I have gained during this interview is actually what I like them to be able to do: [laughs] is to do these questions! Right? And so that's what I'd like to build my case around and then kind of work backwards from there. Like what—what are the pieces that you really need in order to do this question? (Interview 1)	1. Reflection on potential changes to practice 2. Insights into practice	Teacher reflection and learning; instructor as assessor	PPL is a reflective assessor	4

Step 5. Mapping and interpreting

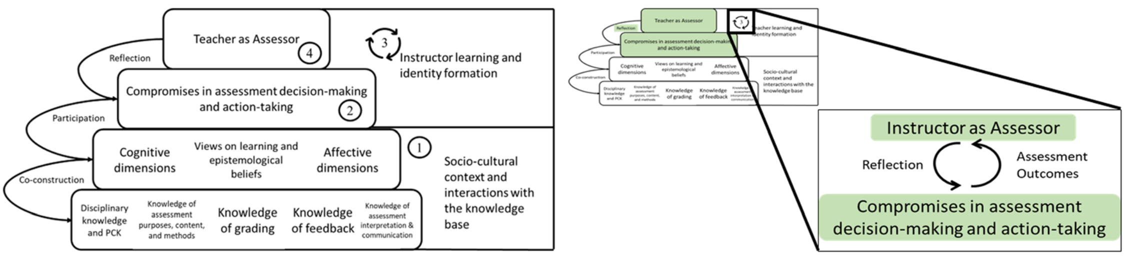
The sub-elements of each of the "levels" of the TALiP were condensed after indexing/coding to provide a way to visualize the overall process. It became apparent over time that there was a dynamic interaction between the levels, and this informed our process of mapping and interpreting the relationship in the data. This process took many iterations and can be broken down into five distinct phases of interpretation and mapping.

Mapping and interpreting pass	Map generated

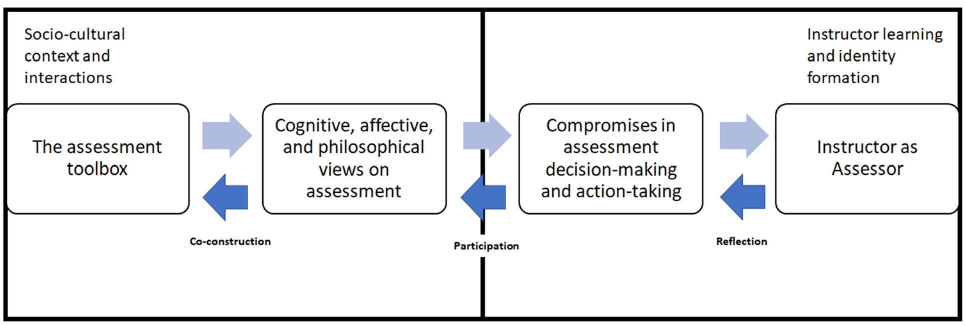
Phase 1



Phase 2



Phase 3



Phase 4

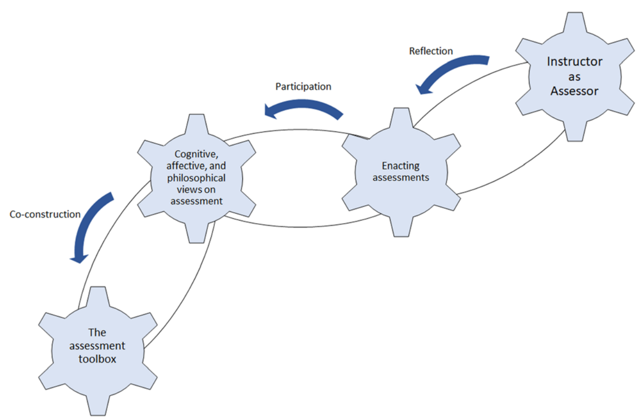
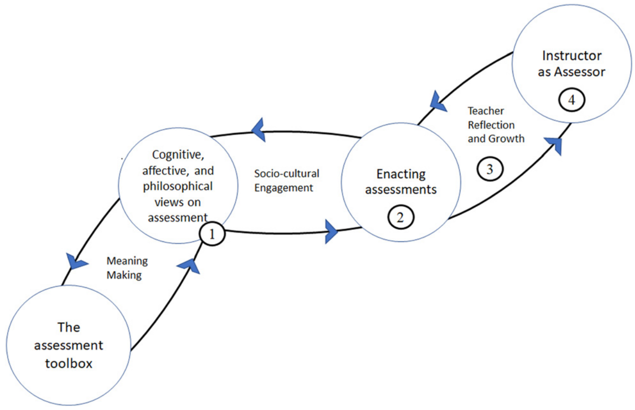


Table (continued)

Mapping and interpreting pass	Map generated
Phase 5 (final)	

Acknowledgements

We would like to acknowledge the support of this work by the National Science Foundation under Grant No. #1224868 and #1849473. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Abell T. N. and Sevian H., (2021), Investigating How Teachers' Formative Assessment Practices Change Across a Year, *J. Chem. Educ.*, **98**(9), 2799–2808, DOI: **10.1021/acs.jchemed.1c00356**.

Adie L., (2013), The development of teacher assessment identity through participation in online moderation, *Assess. Educ.: Princ., Pol. Pract.*, **20**(1), 91–106.

Adie L., Mergler A., Alford J., Chandra V. and Hepple E., (2017), Teacher educators' critical reflection on becoming and belonging to a community of practice, *Communities of Practice*, Springer, pp. 403–419.

Allal L., (2013), Teachers' professional judgment in assessment: A cognitive act and a socially situated practice, *Assess. Educ.: Princ., Pol. Pract.*, **20**(1), 20–34, DOI: **10.1080/0969594X.2012.736364**.

Austin R. C. and Murray T. A., (2019), Don't Go It Alone: The Importance of Community and Research in Implementing and Maintaining Innovative Pedagogy, *Biochemistry Education: From Theory to Practice*, ACS Publications, pp. 247–259.

Ball A. F. and Ladson-Billings G., (2020), Educating teachers for the 21st century: Culture, reflection, and learning, *Handbook of the cultural foundations of learning*, Routledge, pp. 387–403.

Beach A. L., Sorcinelli M. D., Austin A. E. and Rivard J. K., (2016), *Faculty development in the age of evidence: Current practices, future imperatives*, Stylus Publishing, LLC.

Bell E., Provost J. and Bell J. K., (2019), Skills and Foundational Concepts for Biochemistry Students, in Bussey T. J., Linenberger Cortes K. and Austin R. C. (ed.), *ACS Symposium Series*, American Chemical Society, vol. 1337, pp. 65–109, DOI: **10.1021/bk-2019-1337.ch004**.

Bouwma-Gearhart J., Perry K. H. and Presley J. B., (2014), Improving Postsecondary STEM Education: Strategies for Successful Interdisciplinary Collaborations and Brokering Engagement With Education Research and Theory, *J. Coll. Sci. Teach.*, **44**(1), 40–47.

Coombs A., DeLuca C., LaPointe-McEwan D. and Chalas A., (2018), Changing approaches to classroom assessment: An empirical study across teacher career stages, *Teach. Teach. Educ.*, **71**, 134–144, DOI: **10.1016/j.tate.2017.12.010**.

Coombs A., DeLuca C. and MacGregor S., (2020), A person-centered analysis of teacher candidates' approaches to assessment, *Teach. Teach. Educ.*, **87**, 102952, DOI: **10.1016/j.tate.2019.102952**.

Cowie B., Cooper B. and Ussher B., (2014), Developing an identity as a teacher-assessor: Three student teacher case studies, *Assess. Mater.*, **7**(6), 64–89.

Cowie B. and Trevethan H., (2021), Funds of Knowledge and Relations as a Curriculum and Assessment Resource in Multicultural Primary Science Classrooms: A Case Study from Aotearoa New Zealand: A case study from Aotearoa New Zealand, in *International Handbook of Research on Multicultural Science Education*, Cham: Springer International Publishing, pp. 1–32.

Crisp G. T., (2012), Integrative assessment: Reframing assessment practice for current and future learning, *Assess. Eval. Higher Educ.*, **37**(1), 33–43, DOI: **10.1080/02602938.2010.494234**.

DeLuca C. and Klinger D. A., (2010), Assessment literacy development: Identifying gaps in teacher candidates' learning, *Assess. Educ.: Princ., Pol. Pract.*, **17**(4), 419–438, DOI: **10.1080/0969594X.2010.516643**.

DeLuca C. and Johnson S., (2017), Developing assessment capable teachers in this age of accountability, *Assess. Educ.: Princ., Pol. Pract.*, **24**(2), 121–126, DOI: **10.1080/0969594X.2017.1297010**.

DeLuca C., Valiquette A., Coombs A., LaPointe-McEwan D. and Luhanga U., (2018), Teachers' approaches to classroom assessment: A large-scale survey, *Assess. Educ.: Princ., Pol. Pract.*, **25**(4), 355–375.

DeLuca C., Coombs A. and LaPointe-McEwan D., (2019), Assessment mindset: Exploring the relationship between teacher mindset and approaches to classroom assessment, *Stud. Educ. Eval.*, **61**, 159–169, DOI: 10.1016/j.stueduc.2019.03.012.

DeLuca C., Rickey N. and Coombs A., (2021), Exploring assessment across cultures: Teachers' approaches to assessment in the US, China, and Canada, *Cogent Educ.*, **8**(1), 1921903.

Demirdogen B. and Korkut H. M., (2021), Does Teacher Education Matter? Comparison of Education and Science Major Teachers' Assessment Literacy, *Eura. J. Educ. Res.*, **21**(26), DOI: 10.14689/enad.26.2.

Deng D., Xu C., Sun P., Wu J., Yan C., Hu M. and Yan N., (2014), Crystal structure of the human glucose transporter GLUT1, *Nature*, **510**(7503), 121–125.

Emenike M., Raker J. R. and Holme T., (2013), Validating Chemistry Faculty Members' Self-Reported Familiarity with Assessment Terminology, *J. Chem. Educ.*, **90**(9), 1130–1136, DOI: 10.1021/ed400094j.

Entwistle N., (2008), Threshold concepts and transformative ways of thinking within research into higher education, *Threshold concepts within the disciplines*, Brill Sense, pp. 21–35.

Esteban-Guitart M. and Moll L. C., (2014), Funds of Identity: A new concept based on the Funds of Knowledge approach, *Cult. Psych.*, **20**(1), 31–48, DOI: 10.1177/1354067X13515934.

Forsberg E. and Wermke W., (2012), Knowledge sources and autonomy: German and Swedish teachers' continuing professional development of assessment knowledge, *Prof. Dev. Educ.*, **38**(5), 741–758, DOI: 10.1080/19415257.2012.694369.

Frey N. and Fisher D., (2009), Using common formative assessments as a source of professional development in an urban American elementary school, *Teach. Teach. Educ.*, **25**(5), 674–680, DOI: 10.1016/j.tate.2008.11.006.

Fulmer G. W., Lee I. C. H. and Tan K. H. K., (2015), Multi-level model of contextual factors and teachers' assessment practices: An integrative review of research, *Assess. Educ.: Princ., Pol. Pract.*, **22**(4), 475–494, DOI: 10.1080/0969594X.2015.1017445.

Gast I., Schildkamp K. and van der Veen J. T., (2017), Team-Based Professional Development Interventions in Higher Education: A Systematic Review, *Rev. Educ. Res.*, **87**(4), 736–767, DOI: 10.3102/0034654317704306.

Gehrke S. and Kezar A., (2017), The roles of STEM faculty communities of practice in institutional and departmental reform in higher education, *Am. Educ. Res. J.*, **54**(5), 803–833.

Gibbons R. E., Reed J. J., Srinivasan S., Murphy K. L. and Raker J. R., (2022), Assessment Tools in Context: Results from a National Survey of Postsecondary Chemistry Faculty, *J. Chem. Educ.*, **99**(8), 2843–2852.

Gillespie K. H., Robertson D. L. and Bergquist W. H., (2010), *A guide to faculty development*, San Francisco: Jossey-Bass.

Goldsmith L. J., (2021), Using Framework Analysis in Applied Qualitative Research, *Qual. Rep.*, **26**(6), 2061–2076, DOI: 10.46743/2160-3715/2021.5011.

González N., Moll L. C. and Amanti C. (ed.), (2006), *Funds of knowledge: Theorizing practices in households, communities, and classrooms*, Routledge.

Green D. A., Loertscher J., Minderhout V. and Lewis J. E., (2017), For want of a better word: Unlocking threshold concepts in natural sciences with a key from the humanities? *Higher Educ. Res. Dev.*, **36**(7), 1401–1417, DOI: 10.1080/07294360.2017.1325848.

Guba E. G. and Lincoln Y. S., (1982), Epistemological and methodological bases of naturalistic inquiry, *ECTJ*, **30**(4), 233–252.

Hadar L. and Brody D., (2010), From isolation to symphonic harmony: Building a professional development community among teacher educators, *Teach. Teach. Educ.*, **26**(8), 1641–1651, DOI: 10.1016/j.tate.2010.06.015.

Harper S. R., (2010), An anti-deficit achievement framework for research on students of color in STEM, *New Direc. Inst. Res.*, **2010**(148), 63–74, DOI: 10.1002/ir.362.

Harshman J. and Yezierski E., (2016), Characterizing high school chemistry teachers' use of assessment data via latent class analysis, *Chem. Educ. Res. Pract.*, **17**(2), 296–308, DOI: 10.1039/C5RP00215J.

Herppich S., Praetorius A.-K., Förster N., Glogger-Frey I., Karst K., Leutner D., Behrmann L., Böhmer M., Ufer S., Klug J., Hetmanek A., Ohle A., Böhmer I., Karing C., Kaiser J. and Südkamp A., (2018), Teachers' assessment competence: Integrating knowledge-, process-, and product-oriented approaches into a competence-oriented conceptual model, *Teach. Teach. Educ.*, **76**, 181–193, DOI: 10.1016/j.tate.2017.12.001.

Herridge M. and Talanquer V., (2021), Dimensions of Variation in Chemistry Instructors' Approaches to the Evaluation and Grading of Student Responses, *J. Chem. Educ.*, **98**, 270–280.

Herridge M., Tashiro J. and Talanquer V., (2021), Variation in chemistry instructors' evaluations of student written responses and its impact on grading, *Chem. Educ. Res. Pract.*, **22**(4), 948–972, DOI: 10.1039/D1RP00061F.

Izci K. and Siegel M. A., (2019), Investigation of an Alternatively Certified New High School Chemistry Teacher's Assessment Literacy, *Int. J. Educ. in Math., Sci. Tech.*, **7**(1), 1–19, DOI: 10.18404/ijemst.473605.

Jacob M. M., Sabzalian L., Jansen J., Tobin T. J., Vincent C. G. and LaChance K. M., (2018), The Gift of Education: How Indigenous Knowledges Can Transform the Future of Public Education, *Int. J. Mult. Educ.*, **20**(1), 157–185, DOI: 10.18251/ijme.v20i1.1534.

Kiernan M. D. and Hill M., (2018), Framework analysis: A whole paradigm approach, *Qual. Res. J.*, **18**(3), 248–261, DOI: 10.1108/QRJ-D-17-00008.

Koh K. H., (2011), Improving teachers' assessment literacy through professional development, *Teach. Educ.*, **22**(3), 255–276, DOI: 10.1080/10476210.2011.593164.

Ladson-Billings G., (1995), Toward a Theory of Culturally Relevant Pedagogy, *Am. Educ. Res. J.*, **32**(3), 465–491, DOI: 10.3102/00028312032003465.

Levy-Vered A. and Nasser-Abu Alhija F., (2018), The power of a basic assessment course in changing preservice teachers' conceptions of assessment, *Stud. Educ. Eval.*, **59**, 84–93.

Lincoln Y. S. and Guba E. G., (1985), *Naturalistic inquiry*, Sage.

Loertscher J., (2011), Threshold concepts in biochemistry, *Biochem. Mol. Bio Educ.*, **39**(1), 56–57, DOI: 10.1002/bmb.20478.

Loertscher J., Green D., Lewis J. E., Lin S. and Minderhout V., (2014), Identification of Threshold Concepts for Biochemistry, *CBE–LSE*, **13**(3), 516–528, DOI: 10.1187/cbe.14-04-0066.

Loertscher J., Lewis J. E., Mercer A. M. and Minderhout V., (2018), Development and use of a construct map framework to support teaching and assessment of noncovalent interactions in a biochemical context, *Chem. Educ. Res. Pract.*, **19**(4), 1151–1165, DOI: 10.1039/C8RP00029H.

Looney A., Cumming J., van Der Kleij F. and Harris K., (2018), Reconceptualising the role of teachers as assessors: Teacher assessment identity, *Assess. Educ.: Princ., Pol. Pract.*, **25**(5), 442–467, DOI: 10.1080/0969594X.2016.1268090.

Lukin L. E., Bandalos D. L., Eckhout T. J. and Mickelson K., (2004), Facilitating the Development of Assessment Literacy, *Educ. Meas.: Iss. Pract.*, **23**(2), 26–32, DOI: 10.1111/j.1745-3992.2004.tb00156.x.

Luxford C. and Holme T., (2016), How Do Chemistry Educators View Items That Test Conceptual Understanding? *Technology and Assessment Strategies for Improving Student Learning in Chemistry*, American Chemical Society, vol. 1235, pp. 195–210, DOI: 10.1021/bk-2016-1235.ch011.

Mansour N., (2009), Science teachers' beliefs and practices: Issues, implications and research agenda, *Int. J. Environ. Sci. Educ.* **4**(1), 25–48.

Massey K. D., DeLuca C. and LaPointe-McEwan D., (2020), Assessment Literacy in College Teaching: Empirical Evidence on the Role and Effectiveness of a Faculty Training Course, *To Imp. Acad.*, **39**(1), 209–238, DOI: 10.3998/tia.17063888.0039.109.

McCarty T. and Lee T., (2014), Critical Culturally Sustaining/Revitalizing Pedagogy and Indigenous Education Sovereignty, *Harv. Educ. Rev.*, **84**(1), 101–124, DOI: 10.17763/haer.84.1.q83746nl5pj34216.

McMillan J. H., (2003), Understanding and Improving Teachers' Classroom Assessment Decision Making: Implications for Theory and Practice, *Educ. Meas.: Iss. Pract.*, **22**(4), 34–43, DOI: 10.1111/j.1745-3992.2003.tb00142.x.

Meyer J. and Land R., (2006), *Overcoming Barriers to Student Understanding: Threshold Concepts and Troublesome Knowledge*, Routledge.

Moll L. C., Amanti C., Neff D. and Gonzalez N., (1992), Funds of Knowledge for Teaching: Using a Qualitative Approach to Connect Homes and Classrooms, *Theor. Pract.*, **31**(2), 132–141.

Montenegro E. and Jankowski N. A., (2017), *Equity and Assessment: Moving towards Culturally Responsive Assessment. Occasional Paper# 29*, National Institute for Learning Outcomes Assessment.

Montenegro E. and Jankowski N. A., (2020), *A New Decade for Assessment: Embedding Equity into Assessment Praxis. Occasional Paper No. 42*, National Institute for Learning Outcomes Assessment.

Murray T. A., Higgins P., Minderhout V. and Loertscher J., (2011), Sustaining the development and implementation of student-centered teaching nationally: The importance of a community of practice, *Biochem. Mol. Bio Educ.*, **39**(6), 405–411, DOI: 10.1002/bmb.20537.

Offerdahl E. G. and Montplaisir L., (2014), Student-generated reading questions: Diagnosing student thinking with diverse formative assessments, *Biochem. Mol. Bio Educ.*, **42**(1), 29–38, DOI: 10.1002/bmb.20757.

Offerdahl E. G. and Tomanek D., (2011), Changes in instructors' assessment thinking related to experimentation with new strategies, *Assess. Eval. Higher Educ.*, **36**(7), 781–795, DOI: 10.1080/02602938.2010.488794.

Paek P. L. and Holme T. A., (2013), Lessons Learned from Collaborations in Chemistry Assessment across Universities: Challenges in Transfer and Scale, in Holme T., Cooper M. M. and Varma-Nelson P. (ed.), *ACS Symposium Series*, American Chemical Society, vol. 1145, pp. 157–169, DOI: 10.1021/bk-2013-1145.ch011.

Parkinson S., Eatough V., Holmes J., Stapley E. and Midgley N., (2016), Framework analysis: A worked example of a study exploring young people's experiences of depression, *Qual. Res. Psych.*, **13**(2), 109–129, DOI: 10.1080/14780887.2015.1119228.

Perkins D., (2008), in Land R., Meyer J. H. F. and Smith J. (ed.), *Threshold concepts within the disciplines*, Sense Publishers.

Potari D., (2021), Mathematics teacher professional learning and teacher education practices, *J. Math. Teach. Educ.*, **24**(3), 227–230, DOI: 10.1007/s10857-021-09501-8.

Raker J. R., Emenike M. E. and Holme T. A., (2013), Using Structural Equation Modeling To Understand Chemistry Faculty Familiarity of Assessment Terminology: Results from a National Survey, *J. Chem. Educ.*, **90**(8), 981–987, DOI: 10.1021/ed300636m.

Raker J. R. and Holme T. A., (2014), Investigating faculty familiarity with assessment terminology by applying cluster analysis to interpret survey data, *J. Chem. Educ.*, **91**(8), 1145–1151.

Reimann N. and Sadler I., (2017), Personal understanding of assessment and the link to assessment practice: The perspectives of higher education staff, *Assess. Eval. Higher Educ.*, **42**(5), 724–736, DOI: 10.1080/02602938.2016.1184225.

Ritchie J. and Spencer L., (2002), Qualitative data analysis for applied policy research, *Analyzing qualitative data*, Routledge, pp. 187–208.

Sadler I., (2012), The influence of interactions with students for the development of new academics as teachers in higher education, *Higher Educ.*, **64**(2), 147–160, DOI: 10.1007/s10734-012-9545-0.

Samuelson C. C. and Litzler E., (2016), Community Cultural Wealth: An Assets-Based Approach to Persistence of Engineering Students of Color, *J. Eng. Educ.*, **105**(1), 93–117, DOI: 10.1002/jee.20110.

Scarino A., (2013), Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning, *Lang. Test.*, **30**(3), 309–327, DOI: 10.1177/0265532213480128.

Schafer A. G. L. and Yezierski E. J., (2020), Chemistry critical friendships: Investigating chemistry-specific discourse within a domain-general discussion of best practices for inquiry assessments, *Chem. Educ. Res. Pract.*, **21**(1), 452–468, DOI: **10.1039/C9RP00245F**.

Schafer A. G., Borland V. M. and Yezierski E. J., (2021), Visualizing chemistry teachers' enacted assessment design practices to better understand barriers to ''best practices'', *Chem. Educ. Res. Pract.*, **22**(2), 457–475.

Schafer A. G. and Yezierski E. J. (2021a), Investigating how assessment design guides high school chemistry teachers' interpretation of student responses to a planned, formative assessment, *J. Chem. Educ.*, **98**(4), 1099–1111.

Schafer A. G. and Yezierski E. J., (2021b), Investigating high school chemistry teachers' assessment item generation processes for a solubility lab, *Chem. Educ. Res. Pract.*, **22**(1), 214–225, DOI: **10.1039/D0RP00121J**.

Schultz M., Lawrie G. A., Bailey C. H. and Dargaville B. L., (2018), Characterisation of teacher professional knowledge and skill through content representations from tertiary chemistry educators, *Chem. Educ. Res. Pract.*, **19**(2), 508–519, DOI: **10.1039/C7RP00251C**.

Superfine A. C., (2021), An asset-based perspective on prospective teacher education, *J. Math. Teach. Educ.*, **24**(4), 331–333, DOI: **10.1007/s10857-021-09503-6**.

Takahashi S., (2011), Co-constructing efficacy: A ''communities of practice'' perspective on teachers' efficacy beliefs, *Teach. Teach. Educ.*, **27**(4), 732–741, DOI: **10.1016/j.tate.2010.12.002**.

Talanquer V., Bolger M. and Tomanek D., (2015), Exploring prospective teachers' assessment practices: Noticing and interpreting student understanding in the assessment of written work, *J. Res. Sci. Teach.*, **52**(5), 585–609, DOI: **10.1002/tea.21209**.

Tansey J. T., Baird Jr. T., Cox M. M., Fox K., Knight J., Sears D. W. and Bell E., (2013), Core Concepts of Biochemistry and Molecular Biology, *FASEB J.*, **27**, 838.12.

Tienson-Tseng H. L., (2019), Best Practices in Summative Assessment, in *Biochemistry Education: From Theory to Practice*, American Chemical Society, pp. 219–243.

Tomkin J. H., Beilstein S. O., Morphew J. W. and Herman G. L., (2019), Evidence that communities of practice are associated with active learning in large STEM lectures, *Int. J. STEM Educ.*, **6**, 1–15.

Villafañe S. M., Bailey C. P., Loertscher J., Minderhout V. and Lewis J. E., (2011a), Development and analysis of an instrument to assess student understanding of foundational concepts before biochemistry coursework, *Biochem. Mol. Bio Educ.*, **39**(2), 102–109, DOI: **10.1002/bmb.20464**.

Villafañe S. M., Loertscher J., Minderhout V. and Lewis J. E., (2011b), Uncovering students' incorrect ideas about foundational concepts for biochemistry, *Chem. Educ. Res. Pract.*, **12**(2), 210–218, DOI: **10.1039/C1RP90026A**.

Villafañe S. M., Heyen B. J., Lewis J. E., Loertscher J., Minderhout V. and Arnold Murray T., (2016), Design and testing of an assessment instrument to measure understanding of protein structure and enzyme inhibition in a new context, *Biochem. Mol. Bio Educ.*, **44**(2), 179–190, DOI: **10.1002/bmb.20931**.

Villafañe S. M., Minderhout V., Heyen B. J., Lewis J. E., Manley A., Murray T. A., Tienson-Tseng H. and Loertscher J., (2021), Design and Implementation of a Tool to Assess Students' Understanding of Metabolic Pathways Dynamics and Regulation, *CBE–LSE*, **20**(3), ar35, DOI: **10.1187/cbe.20-04-0078**.

Wang J., (2020), Exploring the Perceived Integrations Between Assessment and Metacognition: A Qualitative Inquiry of Three Award-Winning Teacher Educators' Conceptions of Assessment in a Hong Kong University Context, *Front. Educ.*, **4**, 157, DOI: **10.3389/feduc.2019.00157**.

Watson L. A., Bentley A. K., Eppley H. J. and Lin S., (2020), Building an Online Community of Practice for the Evolution of Effective, Evidence-Based Teaching Practices: 15 Years of Improving Inorganic Chemistry Education, *Advances in Teaching Inorganic Chemistry Volume 2: Laboratory Enrichment and Faculty Community*, American Chemical Society, vol. 1371, pp. 127–142, DOI: **10.1021/bk-2020-1371.ch011**.

Wenger E., (1999), *Communities of practice: Learning, meaning, and identity*, Cambridge University Press.

White H. B., Benore M. A., Sumter T. F., Caldwell B. D. and Bell E., (2013), What skills should students of undergraduate biochemistry and molecular biology programs have upon graduation? *Biochem. Mol. Bio Educ.*, **41**(5), 297–301, DOI: **10.1002/bmb.20729**.

Wyatt-Smith C., Klenowski V. and Gunn S., (2010), The centrality of teachers' judgement practice in assessment: A study of standards in moderation, *Assess. Educ.: Princ., Pol. Pract.*, **17**(1), 59–75, DOI: **10.1080/09695940903565610**.

Xu X., Lewis J. E., Loertscher J., Minderhout V. and Tienson H. L., (2017), Small Changes: Using Assessment to Direct Instructional Practices in Large-Enrollment Biochemistry Courses, *CBE–LSE*, **16**(1), ar7, DOI: **10.1187/cbe.16-06-0191**.

Xu Y. and Brown G. T. L., (2016), Teacher assessment literacy in practice: A reconceptualization, *Teach. Teach. Educ.*, **58**, 149–162, DOI: **10.1016/j.tate.2016.05.010**.

Yan X., Zhang C. and Fan J. J., (2018), "Assessment knowledge is important, but...": How contextual and experiential factors mediate assessment practice and training needs of language teachers, *System*, **74**, 158–168, DOI: **10.1016/j.system.2018.03.003**.