He et al. Discip Interdscip Sci Educ Res (2023) 5:22
https://doi.org/10.1186/s43031-023-00088-z

Disciplinary and Interdisciplinary
Science Education Research

RESEARCH Open Access

Transforming standards into classrooms for knowledge-in-use: an effective and coherent project-based learning system

Peng He[1*], Joseph Krajcik[1] and Barbara Schneider[2]

Abstract

Global science education reform calls for developing student knowledge-in-use that applies the integrated knowledge of core ideas and scientific practices to make sense of phenomena or solve problems. Knowledge-in-use development requires a long-term, standards-aligned, coherent learning system, including curriculum and instruction, assessment, and professional learning. This paper addresses the challenge of transforming standards into classrooms for knowledge-in-use and presents an iterative design process for developing a coherent and standards-aligned learning system. Using a project-based learning approach, we present a theory-driven, empirically validated learning system aligned with the U.S. science standards, consisting of four consecutive curriculum and instruction materials, assessments, and professional learning to support students' knowledge-in-use in high school chemistry. We also present the iterative development and testing process with empirical evidence to support the effectiveness of our learning system in a five-year NSF-funded research project. This paper discusses the theoretical perspectives of developing an NGSS-aligned, coherent, and effective learning system and recaps the development and testing process by unpacking all essential components in our learning system. We conclude that our theory-driven and empirically validated learning system would inform high school teachers and researchers across countries in transforming their local science standards into curriculum materials to support students' knowledge-in-use development.

Keywords Science standards, Learning system, Project-based learning, High school chemistry, Student engagement, Curriculum, Instruction, Assessment, Professional learning

Introduction

In a rapidly changing era, an urgent need exists for competitive workers and citizens who can make informed decisions (National Research Council [NRC], 2011; Organization for Economic Cooperation and Development [OECD], 2019; People's Republic of China Ministry of Education, 2014). Science education responds to this need by calling for science education researchers to develop curriculum materials that support learners' knowledge-in-use (aka *usable knowledge*, NRC, 2000) in making informed decisions (Chinese Ministry of Education, 2017; Next Generation Science Standards [NGSS] Lead States, 2013; OECD, 2019; Pellegrino & Hilton, 2012). Knowledge-in-use emphasizes that science teaching needs to support all learners, not to memorize rote facts but to apply their knowledge in new and challenging situations. (He et al., 2022; NRC, 2006; Pellegrino & Hilton, 2012). To meet this vision, nations worldwide need to design curriculum and assessment materials and facilitate teacher professional learning to support knowledge-in-use (Chinese Ministry of Education, 2017; Finnish National Board of Education, 2016; NGSS, Lead States,

*Correspondence:
Peng He
hepeng1@msu.edu
[1] CREATE for STEM Institute, Michigan State University, East Lansing, MI 48824, USA
[2] College of Education, Michigan State University, East Lansing, MI 48824, USA

2013; OECD, 2019). For instance, the Finnish science curriculum addresses knowledge-in-use with several learning competencies for students, such as designing and evaluating scientific inquiry and using scientific concepts (Inkinen et al., 2020). In the U.S., policy documents, such as *the Framework for K-12 Science Education* (2012, hereafter the *Framework*), emphasize knowledge-in-use as students actively engage in science and engineering practices to make sense of real-world phenomena and solve problems by applying crosscutting concepts and disciplinary core ideas, known as three-dimensional (3D) learning. However, teachers cannot develop students' knowledge-in-use over a short period. Developing knowledge-in-use requires comprehensive and systematic support over time. More importantly, students' engagement and achievement in science in the United States have stagnated over the past two decades (National Center for Education Statistics, 2017). In the U.S., the *Framework* (NRC, 2012) and the Next Generation Science Standards (NGSS; NGSS Lead States, 2013) highlight progressively integrating 3D learning of science across K-12 science education to build students' proficiency and appreciation for science over multiple years of school. As such, designing a coherent and effective standards-aligned learning system to support knowledge-in-use has received extensive attention (Fulmer et al., 2018; Roseman et al., 2010; Schmidt et al., 2005). However, designing, implementing, and testing such learning systems is challenging. To date, relatively few empirical studies have investigated the development of a long-term standards-aligned learning system with a series of coherent and consecutive units to show the effectiveness of the learning system in prompting students' academic achievement and motivation.

Project-based learning (PBL) has received extensive attention in science education across several decades (e.g., Blumenfeld et al., 1991; Condliffe, 2017; Haas et al., 2021; Zhao & Wang, 2022). PBL offers a coherent design approach that enables students to immerse in authentic questions to collaboratively design physical artifacts (e.g., Krajcik & Czerniak, 2018). In a PBL learning environment, students develop the ability to apply knowledge over time to mutually reinforce ideas, scientific practices, and problem-solving capacity (Li, Miller, & Krajcik, 2023; Reiser, 2014). Research shows that students who experienced PBL as the instructional focus obtain greater science achievement than those who did not experience such PBL instruction (e.g., Geier et al., 2008; Krajcik et al., 2023; Schneider et al., 2022). Students also demonstrated high-level engagement in PBL environments (e.g., Inkinen et al., 2020; Schneider et al., 2016). High school science is a gatekeeper for many specializations and postsecondary schooling (Hinojosa et al., 2016). However, most PBL studies are at elementary or middle school levels, and relatively few are conducted at the high school level. To address the above challenge, we present an iterative design process to develop a coherent and effective standards-aligned learning system to support high school students' engagement and knowledge-in-use development in a collaborative project titled *Crafting Engagement in Science Environments* (hereafter *CESE*, Schneider et al., 2020). This paper first discusses the conceptual framework of our learning system. It provides an overview of our learning system, design framework, and research phases of developing and testing our learning system. Through a retrospective review of our design process, we unpack the iterative design process of our learning system, including teacher and student curriculum materials, assessment, and professional learning, to support high school students' knowledge-in-use.

Conceptual framework

We design our *CESE* learning system using four design principles: three-dimensional learning (NRC, 2012), PBL instruction (Krajcik & Czerniak, 2018), student-situated engagement (Schneider et al., 2016; Schmidt et al., 2018), and coherent progressive learning supports (Duschl et al., 2011; He et al., in press, a). Figure 1 presents the conceptual framework of the *CESE* learning system.

Developing knowledge-in-use through three-dimensional learning

Knowledge-in-use requires students to investigate the world as scientists, applying scientific core ideas, scientific practices, and crosscutting concepts to solve complex problems or explain compelling phenomena (NRC, 2012). The *Framework* and the NGSS advance 3D learning as a critical avenue to achieve the goal of students developing knowledge-in-use. The NGSS adopted the three dimensions from the Framework (2012) and articulated a series of performance expectations (PEs) to demonstrate student achievement on knowledge-in-use learning goals at the end of each grade band. However, the scope of PEs is too broad at an operational level for teaching and assessment, which requires finer-grain performance learning goals, called learning performances, to explicitly support knowledge-in-use development (Harris et al., 2019; He, Chen, et al., 2023; Li, Chen, et al., 2023). Learning performances retain the structure of 3D knowledge but are smaller in scope than the NGSS PE (Harris et al., 2019; He, Zhai, et al., 2023). A series of 3D learning performances build toward each other to cover the targeting NGSS PE(s) and make them teachable and assessable. Moreover, these fine-grain size learning performances are the learning goals for developing 3D learning activities in curriculum and instructional

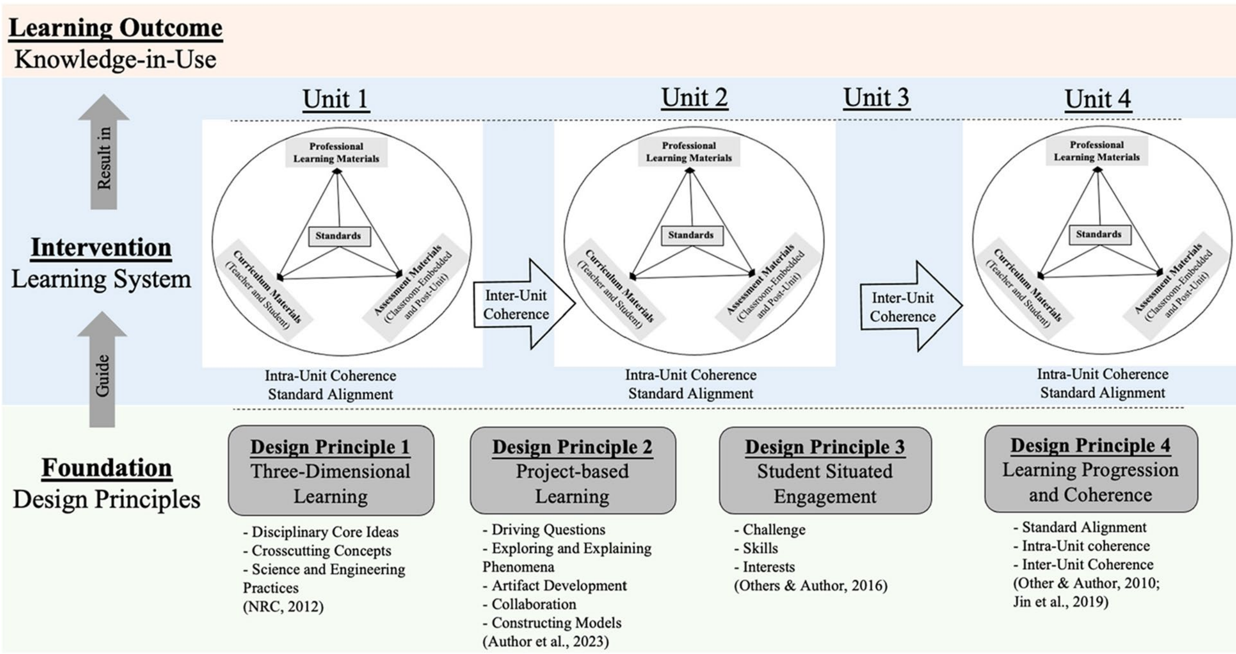


Fig. 1 The Conceptual Framework

materials and assessment tasks to advance student knowledge-in-use.

Developing knowledge-in-use through project-based learning

The PBL approach has been recognized as an efficient vehicle to support 3D learning that results in knowledge-in-use proficiency (Haas et al., 2021; He, Chen, et al., 2023; Li, Chen, et al., 2023; Zhao & Wang, 2022). PBL has three main theoretical underpinnings, including situated cognition theory (e.g., Brown et al., 1989), social constructivism theory (Vygotsky, 1978), and active construction theory (Bransford et al., 2000). As an inquiry-based approach, PBL allows students to make sense of phenomena or design solutions to problems using the 3-dimensions of scientific knowledge (Haas et al., 2021; Li, Miller, & Krajcik, 2023; Reiser et al., 2021). Reviewing the existing literature, we found that curriculum developers and teachers use a principled design process to establish NGSS-aligned PBL instructions (e.g., He, Chen, et al., 2023; Miller & Krajcik, 2019; Reiser et al., 2021). Those previous studies have shared a consensus on the design process, including a) selecting, inspecting, and unpacking the PEs; b) selecting phenomena and driving questions; c) developing lesson-level learning performance goals; d) constructing unit sequences and lessons; and finally, e) assessing student learning outcomes. Moreover, six essential features are necessary for designing a high-quality PBL learning system (Krajcik & Czerniak, 2018): 1) driving students' experience using an anchoring phenomenon linked to driving questions, 2) targeting learning performance goals, 3) engaging in scientific practices, 4) participating in collaborative activities, 5) scaffolding student learning, and 6) creating a set of tangible products.

Developing knowledge-in-use through situated engagement

Knowledge-in-use is grounded in the theory of situated cognition (Brown et al., 1989). Situated cognition recognizes that real-world situations make learning meaningful for learners to make sense of phenomena and motivate them to solve problems in unfamiliar contexts. Students engage in a sensemaking process that uses different scientific practices in diverse contexts that align with the nature of knowledge-in-use. Situational engagement occurs when students experience high levels of challenge, skill, and interest necessary to obtain optimal learning moments (Schneider et al., 2016; Schmidt et al., 2018). The research shows that teacher-facing materials can support students in reaching optimal learning moments (Schneider et al., 2016). Designers can develop optimal learning environments (He, Chen, et al., 2023). by 1) constructing meaningful driving questions that intrigue students and engage them in making sense of the phenomena or problems relevant to the driving questions; 2) allowing learners to investigate meaningful driving questions aligned to fine-grain size 3D learning performance goals; 3) providing opportunities for learners to develop artifacts through individual or collaborative activities, and 4) providing time so learners can endeavor to figure out the driving questions.

Developing knowledge-in-use through coherent progressive student learning

Students' knowledge-in-use development needs a coherent and progressive learning system consisting of curriculum materials, instruction, and assessment. Coherent curriculum materials with aligned assessment tasks and professional learning are critical for developing students' knowledge-in-use (e.g., He, Chen, et al., 2023; Li, Chen, et al., 2023). Learning progressions (LPs) are "descriptions of the successively more sophisticated ways of thinking about a topic that can follow one another as children learn about and investigate a topic over a broad span of time" (NRC, 2007, p214). LPs have been proposed as a coherent framework to align curriculum, instruction, and assessment (Duschl et al., 2011; Fortus & Krajcik, 2012; He et al., in press, a; Jin et al., 2019). Coherence is an essential feature of a learning system that effectively supports students' science learning (NRC, 2006). Fortus and Krajcik (2012) suggest that using LPs can guide different types of coherence in developing curriculum materials: learning goals coherence, intra-unit coherence, and inter-unit coherence. LPs provide a coherent principle for guiding the order of units and the lesson sequence within a unit and for developing learning goals aligned with the standards (He et al., in press, a). Jin et al. (2019) also argue that LPs can provide three types of coherence: (a) the developmental coherence of productive learning across time; (b) horizontal coherence across the curriculum, instruction, and assessment; and (c) vertical coherence between classroom assessments and large-scale assessments. The *CESE* design approach employs coherence principles (Fortus & Krajcik, 2012; Jin et al., 2019) to ensure our learning system's intra-unit and inter-unit coherence.

Our team employed the four design principles as the theoretical foundation. This paper presents a systematic design framework and unpacks the process for iteratively designing a coherent and standards-aligned learning system of teacher-facing and student materials, assessment, and professional learning to support long-term knowledge-in-use development.

Methods

This paper employs design-based research (e.g., Barab & Squire, 2004) as the methodology for developing, implementing, testing, and revising our *CESE* learning system. This section overviews our *CESE* project for our chemistry learning system, introduces the six-stage design framework, summarizes the research phases, and highlights the main takeaways from empirical studies.

Project overview

The *CESE* project was a five-year NSF-funded research project and a collaborative effort among researchers and teachers in the U.S. and Finland (Schneider et al., 2020). The project aimed to increase student engagement and interest in science and support their knowledge-in-use development by designing and testing a PBL learning system. The project also provided ongoing professional learning to help teachers understand the principles of PBL and the use of the *CESE* materials.

Based on the conceptual framework (See Fig. 1), we developed the *CESE* chemistry learning system, consisting of 4 units with teacher-facing and student materials, assessment tasks, and professional learning materials. Our development team created the four chemistry units to support students' knowledge-in-use aligned with the disciplinary core ideas of *Matter and Its Interaction* (PS1) and related *Energy* (PS3) in the NGSS. The 4 units build on each other based on the inter-unit coherence principles. We provided professional learning sessions to support teacher implementation of those units. We also tracked and monitored student learning over time using student artifacts in classrooms and post-unit assessments and evaluated student achievement using an end-of-year summative test designed by the Michigan Department of Education.

Design framework

To design NGSS-aligned PBL learning systems, we employed the construct-centered design (CCD) approach (Shin et al., 2010) that incorporates the learning-goal-driven design (Krajcik et al., 2008) and evidence-centered design process (ECD, Mislevy & Haertel, 2006). The learning-goal-driven design process ensures that the articulated learning goals for class activities and assessment practices at the lesson or unit levels align with the targeted standards. The learning-goal-driven design process includes unpacking related content standards (e.g., 3D knowledge) and reconstructing a series of grain-size learning goals that can reflect the requirements in the standards. While using the ECD approach, we aim to develop curriculum and assessment materials with the evidentiary base for specifying their coherent and logical relationships with the articulated learning goals. Evidence of student performances collected from classroom activities and assessment tasks should reflect the target constructs described in the learning goals.

Accordingly, we articulated a principled six-stage systematic design process (see Fig. 2) for developing NGSS-aligned learning systems. Stage 1 identifies and

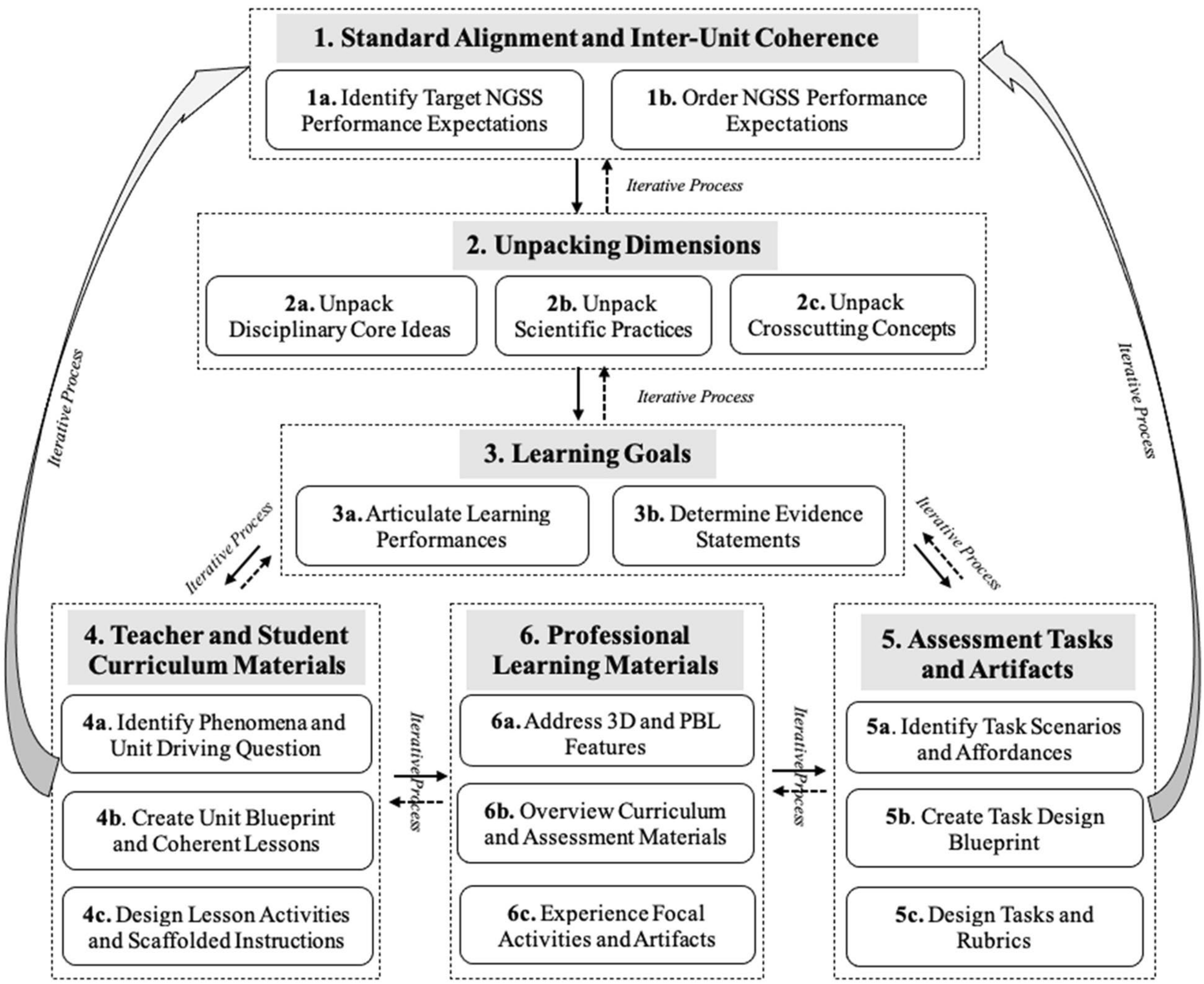


Fig. 2 The Design Framework of Developing NGSS-Aligned PBL Learning System

bundles related performance expectations/goals for developing single units (1a) and orders the bundled PEs into a sequence to ensure inter-unit coherence of curriculum materials (1b). We employed the notion of learning progression as a design principle to guide the intra-unit coherence and inter-unit coherence of the *CESE* learning system with four consecutive units. Stage 2 unpacks the 3D knowledge from the targeted PEs. Unpacking helps designers identify and elaborate on all aspects incorporated in the PEs to demonstrate student knowledge-in-use successfully. The information from the unpacked documents of three dimensions is used to articulate a set of lesson-level and task-specific learning performances and associated evidence statements (Stage 3). Incorporating PBL features (Krajcik & Czerniak, 2018), Stage 4 develops teacher and student curriculum materials, consisting of 4a) identifying the overarching phenomenon and the unit driving questions, 4b) creating a unit blueprint with a coherent lesson sequence; 4c) designing lesson activities and scaffolded instructions. Stage 5 develops assessment tasks and artifacts for classroom use and post-unit assessments. The development process starts with 5a) identifying task scenarios, 5b) creating task design blueprints, and 5c) designing tasks and rubrics. The *CESE* team then developed unit-specific professional learning materials (Stage 6), consisting of 6a) addressing 3D and PBL features, 6b) overviewing curriculum and assessment materials, and 6c) experiencing focal activities and artifacts.

The *CESE* systematic design process is an iterative approach that allows designers to revise and improve the materials during each phase. Moreover, the construct-centered design approach ensures the alignment among standards, curriculum, assessment, instruction, and professional learning in the learning system.

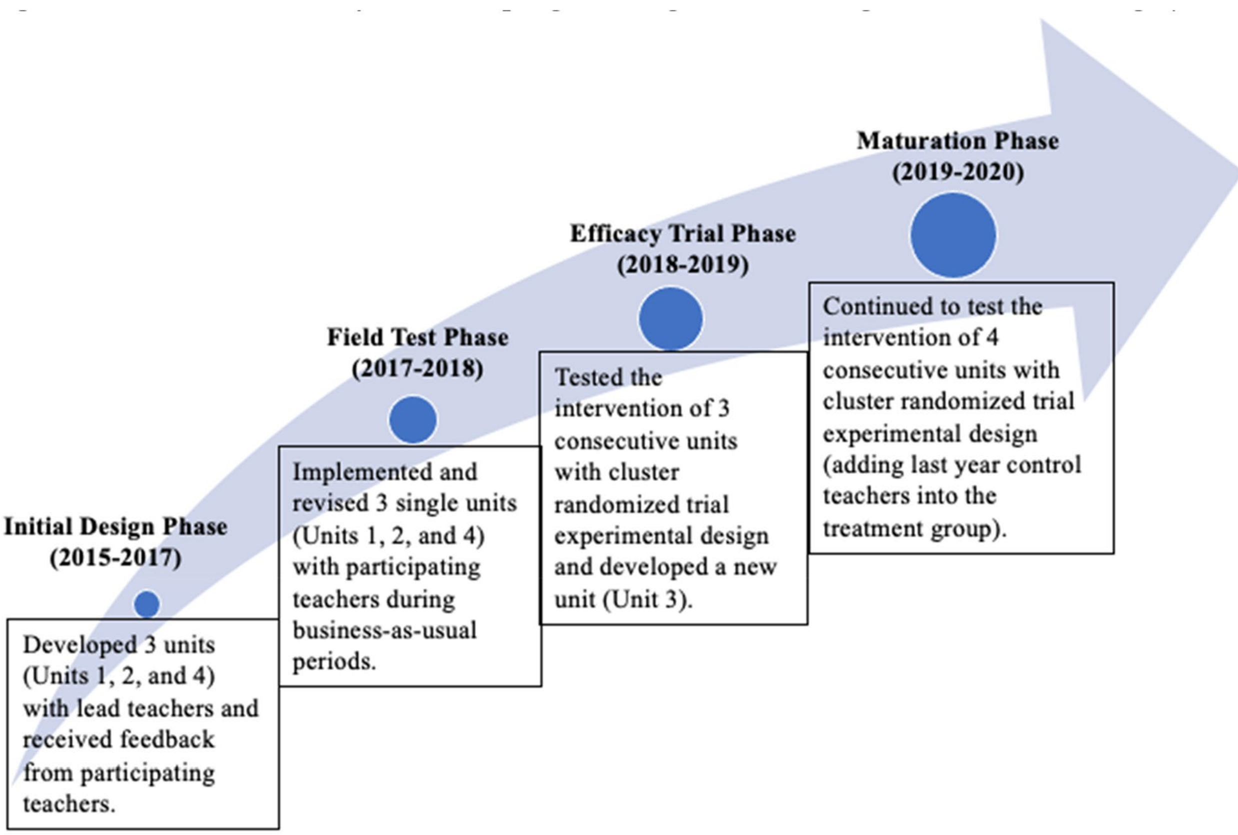


Fig. 3 Research Phases for Developing, Testing, and Revising the CESE Learning System

Research design and Main takeaways

Figure 3 presents the research design, including the initial development phase (2015–2017), field test (2017–2018), and efficacy trial (2018–2019), and maturation study (2019–2020). We elaborate on each phase's main activities and takeaways from empirical research.

In the developmental phase (2015–2016), our curriculum development team co-designed the two chemistry units (the *Evaporative Cooling* unit and the *Conservation of Matter and Atoms* unit) with two lead teachers and worked with five teachers to pilot and revise the initial units (Schneider et al., 2020). In 2016–2017, our team continually developed the third unit-*Periodical Table*, with a lead teacher and piloted it with five new participating teachers. Our co-design and piloting teachers come from the urban, suburban, and rural school districts in a mid-western U.S. state. Based on a qualitative study of teachers' implementation of the units, the prior study (Bielik et al., 2022) from the *CESE* team found that the units helped students build toward the NGSS learning goals; the unit driving questions were critical in providing lessons with coherence and relevance; students were meaningfully using scientific practices, mostly the scientific practice of developing and using scientific modeling; and students' artifacts contributed to their learning and to the assessment of their learning.

In the field test phase (2017–2018), our research team conducted a single case design with eight participating teachers to implement and revise our curriculum and assessment materials and investigate how our PBL units increase situational engagement using the experience sampling method (ESM, Csikszentmihalyi & Schneider, 2000). Our prior studies from the *CESE* project provide evidence to support student engagement in the PBL learning system using an experience sampling method approach. For instance, Schneider et al. (2016) showed that students are more likely to feel confident, successful, and happy when challenged in science classrooms. In another study, Inkinen et al. (2019) investigated U.S. and Finish students and found that situational engagement is associated with specific science classroom activities, such as analyzing data, constructing models, and presenting scientific information. In addition, Inkinen et al. (2019) also found that two scientific practices (i.e., developing models and constructing explanations) were associated with higher student situational engagement compared to other practices (e.g., asking questions), indicating that the *CESE* learning system prompts students' attention

Table 1 The Overview of Targeting NGSS PEs in the Four Units

Standard Alignment	Unit 1: Evaporative Cooling	Unit 2: Periodic Table	Unit 3: Combustion Reactions	Unit 4: Conservation of Mass and Atom
PS1: Matter and Its Interaction	HS-PS1–3: Plan and conduct an investigation to gather evidence to compare the structure of substances at the bulk scale to infer the strength of electrical forces between particles.	HS-PS1–1: Use the periodic table as a model to predict the relative properties of elements based on the patterns of electrons in the outermost energy level of atoms. HS-PS1–2: Construct and revise an explanation for the outcome of a simple chemical reaction based on the outermost electron states of atoms, trends in the periodic table, and knowledge of the patterns of chemical properties.	HS-PS1–4: Develop a model to illustrate that the release or absorption of energy from a chemical reaction system depends upon the changes in total bond energy.	HS-PS1–2: Construct and revise an explanation for the outcome of a simple chemical reaction based on the outermost electron states of atoms, trends in the periodic table, and knowledge of the patterns of chemical properties. HS-PS1–7: Use mathematical representations to support the claim that atoms, and therefore mass, are conserved during a chemical reaction.
PS3. Energy	HS-PS3–2: Develop and use models to illustrate that energy at the macroscopic scale can be accounted for as a combination of energy associated with the motions of particles (objects) and energy associated with the relative positions of particles.			

in science classrooms through those focal scientific practices.

Following the field study, the team conducted an efficacy study in the United States from 2018 to 2019 (Schneider et al., 2022). The efficacy study involved a diverse group of over 4000 students in a randomized control trial conducted in California and Michigan that tested the effectiveness of the intervention (of the three initial units, Units 1, 2, and 4, see the descriptions in Table 1) on students' science achievement. Sixty-one schools of 119 teachers and 4238 students were randomly assigned to treatment (30 schools, 2127 students) and control groups (31 schools, 2111 students). Teachers and students in the treatment group implemented our curriculum and assessment materials. Teachers received professional learning on NGSS and 3D learning, an overview of our learning system, features of PBL, and unit-specific activities. Meanwhile, teachers and students in the control group used their local curriculum units and only received professional learning to understand the features of NGSS and 3D learning. The efficacy study (2018–2019) results show that students in the intervention of the three consecutive units (Units 1, 2, and 4) performed significantly and substantially higher on the third-party assessment tasks than those in the matched control group (Schneider et al., 2022). We further explored the intervention process and investigated students' 3D learning across our learning system. We found that students' performance on the three post-unit assessments could cumulatively and individually impact their summative science achievement (He, Chen, et al., 2023). The above research provided robust empirical evidence to support the effectiveness of the *CESE* learning system in developing high school students' knowledge-in-use.

In 2019–2020, we continued to test our intervention of the 4 units by adding the control teachers from last year to the treatment group. We also provided professional learning for treatment teachers before they implemented our units. Unfortunately, we could not complete the maturation study due to the Covid-19 outbreak. Most teachers could only teach the first and second units, and only a few teachers completed the third unit. We used the remaining time to interview our teachers regarding how they used the units remotely.

Based on the five-year investigations, the *CESE* team has gathered robust and sufficient evidence to show the learning system's effectiveness in supporting student knowledge-in-use development. These positive results motivate us to unpack the iterative process of developing and revising our effective, coherent, and standards-aligned learning system.

Unpacking the *CESE* chemistry learning system

In this section, we unpack the iterative process of developing our learning system and discuss how our learning system supports students' knowledge-in-use. Accordingly, we unpacked the six-stage design process (see Fig. 2) and specified the development and revision process for designing the materials.

Standards alignment and inter-unit coherence

Identifying and Bundling Target NGSS PEs to Ensure Standards Alignment. To align with the NGSS, *the CESE* team selected the PEs from *Matter and Its Interaction* (HS-PS1) and *Energy* (HS-PS3) to design our high school chemistry learning system (Stage 1a). Our team identified the PEs for the chemistry learning system because they are vital for students to make sense of phenomena related to the structure and properties of matter and interactions of matter and energy (NRC, 2012). During the initial development phase, the *CESE* team developed 3 units: Unit 1 (Evaporative Cooling), Unit 2 (Periodic Table), and Unit 4 (Conservation of Mass and Atoms). Those units were designed primarily based on the PEs related to the *Matter and Its Interactions*. The Evaporative Cooling unit also included a PE (i.e., HS-PS3–2[1]) related to *Energy*. In the field test phase, the *CESE* team modified the bundle of PEs to overlap with the energy PE (i.e., HS-PS3–2) for each unit, enhancing the connections among units. The HS-PS3–2 is an energy PE that addresses how the energy transfer relates to the changes in particles' motions and positions. Particles refer to abstract microscopic objects – atoms, molecules, and outermost electrons and nucleus, which make the HS-PS3–2 essential for connecting multiples PEs (e.g., HS-PS1–3, 1–1, and 1–7) in the *Matter and its Interactions*. In the efficacy study, we found that even though all 3 units (Units 1, 2, and 4) could individually contribute to students' achievement, no significant differences in the single effects between Unit 2 and Unit 4 exist, indicating students might not transfer their understanding in Unit 2 to learning ideas in Unit 4 (He, Chen, et al., 2023). We also received feedback from teachers that they usually added lessons related to the bond energy in chemical reactions between Units 2 and 4. Based on the empirical evidence, we designed a new unit (Combustion Reaction, Unit 3) and placed it between Unit 2 (Periodic Table) and Unit 4 (Conservation of Mass and Atoms) to strengthen the entire learning system (see the list of units in Table 1).

[1] For HS-PS3–2 in the NGSS, HS represents the grade band of High School; PS represents the domain of Physical Sciences; and 3–2 represents the 2nd performance expectation in the 3rd disciplinary core ideas (i.e., Energy).

Ordering a Series of NGSS PEs to Ensure Inter-Unit Coherence. The *CESE* team had not ordered the units before the efficacy study because of conducting a field study of single units. In the development phase, teachers implemented the single units during business-as-usual instructional periods. However, teachers might not implement the units in the same order as different school districts and teachers have their planned business-as-usual schedules, which increased the variations in implementing our units. To successfully implement the unit, teachers usually adjust the lesson plans in the unit because some prior knowledge or learning experiences are required for their students. Starting from the efficacy study, the CESE team ordered the 3 units (Units 1, 2, and 4, see Stage 1b) based on the inter-unit coherence design principle (Fortus & Krajcik, 2012). We specified prerequisite and prior knowledge to connect our units in each unit material. As described above, we added one additional unit (Unit 3) between Unit 2 and Unit 4 to enhance the inter-unit coherence of the CESE learning system. The *CESE* chemistry learning system supports students in explaining temperature and phase change of matter (Unit 1), energy and intra-atom interactions for explaining the properties of elements in the periodic table (Unit 2), energy and intramolecular interactions related to chemical reactions (Unit 3), and conservation of matter and atoms (Unit 4). From theoretical perspectives, we suggest the sequence of the four consecutive and coherent units could support students' progressive knowledge-in-use in matter and related energy at the high school level.

Unpacking dimensions

The *CESE* team unpacked the disciplinary core ideas, crosscutting concepts, and scientific practices for each PE (Stage 2). As each PE encompasses a vast domain of complex constructs, it is impossible to address them in a single lesson. In the unpacking process, the *CESE* team first elaborated on the meanings of the 3D knowledge and identified the sub-aspects in each dimension. For instance, The NGSS PE of HS-PS1–3 states, *"Plan and conduct an investigation to gather evidence to compare the structure of substances at the bulk scale to infer the strength of electrical forces between particles."* Using the unpacking resources (e.g., Duncan et al., 2017), we elaborated the core ideas in this PE: the structure of substances, the relationships between the properties and structures of substances, and the strength of electrical forces between particles. In addition, we identified the core ideas related to this PE at the middle school level to understand students' prior knowledge, such as the properties of substances and the types and arrangement of atoms to form different molecules.

Next, based on existing literature (e.g., Schwarz et al., 2017), we elaborated on the meaning of *Planning and Conducting Investigations*. First, a planned investigation should describe the procedure for collecting data about a substance's properties and elaborate on why the data can be used as evidence to understand the strength of the electrical forces between the particles of the substance. Second, carrying out an investigation should include collecting and recording data (both quantitative and/or qualitative) about the bulk properties of substances and refining the design (the accuracy and precision of the data collected and limitations of the investigation). We also identified several related scientific practices, such *as Developing and Using Models* and *Constructing Scientific Explanations*. Like unpacking the science practice, we elaborated on the meaning of the crosscutting concept of *Patterns* to describe the different data patterns about properties of substances related to the strength of electrical forces between particles (see the reference book, Nordine & Lee, 2021). We further identified several related crosscutting concepts, such as *Structure and Function* and *Cause and Effect*. The unpacking process ensures that the broad and vague statements in the NGSS PEs translate into specific and essential statements of 3D knowledge. During the design process, our design team returned to the unpacking documents to check the coherence between the standards and the following design work (e.g., learning goals).

Learning goals

Through unpacking, we elaborated on the three dimensions of a PE into several specific aspects that we used to articulate a series of learning performance goals and the associate evidence statements (See Stage 3 in Fig. 2). Learning performance goals are 3D statements incorporating aspects of DCIs, SEPs, and CCCs but are smaller in scope than a PE (Harris et al., 2019). A related set of learning performances describes the performance needed to build toward a PE or a bundle of PEs. As learning performance goals are the expected statements that students demonstrated, the associate evidence statements are the observable evidence students must provide to meet the goals. In our study, we articulated a set of learning performance goals and associated evidence statements that we used to develop lessons in the units and assessment tasks. For instance, we constructed 10 learning performances based on unpacking the PE bundle (HS-PS1–3 and 3–2) for the *Evaporative Cooling* unit (i.e., Unit 1). One of the learning performance goals is "*Students develop and use a model that shows that the energy is transferred to the water that causes the water molecules to move faster, resulting in the temperature increases.*" When designing a lesson or an assessment task, our design team contextualized this

Table 2 The Overview of the Four Teacher and Student Unit Materials

	Unit 1: Evaporative cooling	Unit 2: Periodical table	Unit 3: Combustion reactions	Unit 4: Conservation of Mass and Atom
Unit length	Ten lessons in 3–4 weeks	Ten lessons in 3–4 weeks	Ten lessons in 3–4 weeks	Nine lessons in 3–4 weeks
Unit driving question	Why do I feel colder when I am wet than when I am dry?	Why can you eat some substances like table salt (NaCl), but their components (Sodium and Chlorine) are toxic?	Why can we burn fuel to keep ourselves warm?	What happened to my substance?
Focal practices and crosscutting concept	Developing and using models and cause and effect	Developing and using models and cause and effect	Developing and using models and cause and effect	Developing and using models and cause and effect
Unit artifacts	Develop a microscope-level model for explaining the changes in temperature and particles during the evaporation process.	Create a periodic table as a model to explain or predict the relative properties of elements and the outcome of a simple chemical reaction.	Develop an atomic and molecular model to explain the bond energy change in the matter before and after a simple chemical reaction.	Develop an atomic and molecular model to explain the conservation of mass and atoms before and after a simple chemical reaction.

learning performance goal with a specific context in a unit. This learning performance goal was for *Lesson 4* in the *Evaporative Cooling* unit. The evidence statement is "looking for students drawn models to show water particles are moving faster when energy was transferred to the water in a liquid phase and the temperature increased." In *Lesson 4*, students are expected to use the collected data from two activities (i.e., adding dye to cold and warm water) to experience phenomena related to the core idea. The evidence statement expects to see students draw models that show that when water molecules are moving faster, the average kinetic energy of water molecules increases. Evidence statements in each lesson provide teachers with the criteria to analyze and diagnose student performance on the lesson-level learning performance goal using students' artifacts or explanations.

Teacher and student curriculum materials

We designed 4 units to align with the selected NGSS PEs (Stage 4). Table 2 presents an overview of the four teacher and student curriculum materials. Each unit consists of 9–10 lessons. Each unit takes approximately 3–4 weeks to complete. Initially, our design team only allocated *2 weeks* (10 one-hour-long lessons) for each unit in the development phase (2015–2017). We learned from our participating teachers, especially inexperienced teachers, that sufficient time is needed to engage students in the experiences planned for the units (Units 1 and 2). We realized that some activities required more time for students to experience, such as carrying out investigations (e.g., chemical reactions need overnight procedures) and developing, evaluating, and revising models, which took more time than initially expected. We extended those lessons longer and made our unit 3–4 weeks in the subsequent efficacy and maturation phases (2018–2020). Using the PBL features (Krajcik & Czerniak, 2018), the design team designed each unit based on the process in Stage 4 (See Fig. 2). To ensure inter-unit coherence, the design team emphasized the focal practice and crosscutting concept of *Developing and Using Models* and *Cause and Effect* across the 4 units. Focusing on building student proficiency in modeling and using the idea of cause and effect gained in the initial units would support their further development in the latter units. The *CESE* team has empirical evidence to support this position based on testing the 3 units (Units 1, 2, and 4) in our efficacy study (He, Chen, et al., 2023). Moreover, the final artifacts in the 4 units build upon each other because they develop the microscope-level (particle or atomic-molecular) models to explain what happens when the matter changes.

Our team developed each unit based on the design process expressed in Stage 4 (see Fig. 2). We first identified the overarching phenomena and the unit-driving question (see Stage 4a). They are essential features of PBL for establishing relevance to students' lives and enhancing the emotional engagement of students to make sense of the question (Li, Miller, & Krajcik, 2023). We employed the criteria of whether the phenomenon and question are compelling and broad enough to allow students to develop a sense of wonder, feasible to design and carry out an investigation, worthwhile as students can engage in making sense of phenomena, contextualization with real-world issues, sustainable (pursue solution over time), and ethical in that no harm occurs to living organisms or the environment (Krajcik & Czerniak, 2018). For instance, the PEs (HS-PS1–3 and 3–2) for Unit 1 emphasize the changes of phase and temperature in substances using the ideas of particle interactions, their motions, and their associated energy transfer. In the initial development phase, our design team chose an everyday phenomenon about water evaporation and asked the unit driving question, "*When I am sitting by a pool, why do I feel colder when I am wet than when I am dry*?" We received positive feedback from participating teachers that this driving question could sustain activities across the unit. However, when we reflected on the unit implementation, we realized that the phenomenon of "sitting by a pool" may restrict students from asking related questions, which may not be suitable for all students. Instead, the phenomena of water evaporation do not necessarily happen in the "sitting by a pool" scenario. Thus, we modified the unit driving question with a broader context: "*Why do I feel colder when I am wet than when I am dry?*" as the unit driving question. The first unit lesson has students share their experience related to the phenomenon and experience a hands-on activity of putting one hand in the water and another hand out of the water. Based on students' prior and hands-on experience, they ask subquestions related to the unit driving question and share them with the class.

We then created a unit blueprint and a lesson sequence consisting of the learning performances, evidence statements, lesson-level driving questions, and the outlines of learning activities (Stage 4b). In Unit 1, we articulated 10 learning performance goals aligned with the target PEs. As shown in Table 3, Unit 1 has 10 lessons. The unit starts with students experiencing the unit driving questions (in *Lesson 1*) and ends with developing and using models to figure out the driving questions (in *Lesson 10*). *Lessons 2* to *9* involve students collecting sufficient evidence to support claims and diving deeper into the ideas of electrical forces between particles, motions of the particles, and associated energy transfer to make sense of the overarching phenomenon in this unit. Although the unit emphasized modeling practices and cause and effect

Table 3 The Lesson Sequence and Associated 3D Learning Performance Goals for Unit 1-Evaporative Cooling Unit

Lesson-level driving questions	Lesson-level 3D learning performance goals
L1: Why do I feel colder when I am wet than when I am dry?	Students observe water evaporation to ask questions and develop initial models related to energy and matter about why substances evaporate using the lens of cause and effect.
L2: What happens to water when it is heated?	Students conduct an experiment and analyze and interpret data about the temperature changes in water in various phases as it is heated.
L3: How does developing bar chart models help understand the temperature changes that occur while heating water in various phases?	Students develop a model based on the data and observations they collected in the previous activity to explain the relationship between the transfer of energy and temperature of the water as it is heated and undergoes phase changes from solid to liquid and from liquid to gas.
L4: What happens to water when the temperature rises?	Students make observations and interpret data of water in the heated liquid phase to make claims supported by evidence that when energy is transferred to water in the liquid phase, the temperature increases because the molecules move faster. Students develop and use a model that shows that the energy is transferred to water, which causes the water molecules to move faster, increasing the temperature.
L5: What happens to water when the temperature does not change?	Students revise particle-level models using collected data to explain why the water temperature does not change when it melts or boils.
L6: How can we create a model to explain why we feel colder when we are wet than when we are dry?	Students revise the initial models to explain that energy is transferred from skin to liquid when water evaporates.
L7: What are the differences between liquids when evaporating? (Part I)	Students plan and conduct investigations to make connections between the evaporation rates of different substances and the temperature changes that occur during evaporation.
L8: What are the differences between liquids when evaporating? (Part II)	Students analyze and interpret data to make connections between the identity of a chemical, the rate of evaporation, and the temperature change that occurs as a liquid evaporates. Students test and revise models to show relationships among the evaporation rates of different liquids.
L9: Why do liquids have different evaporation rates?	Students explain how differences in the strength of attractive forces between particles can account for different macroscopic properties.
L10: Why do I feel colder when I am wet than when I am dry?	Students revise and use their consensus models to explain the process of evaporative cooling, connecting the energy changes to changes in the structure of matter in the system.

as the focal scientific practice and crosscutting concept, the design team deliberately created 3D learning performance goals with other essential practices and crosscutting concepts, such as "*Analyzing and Interpreting Data with Patterns*" (*Lesson 8*) and "*Constructing Scientific Explanations with Cause and Effect*" (*Lesson 9*). Across the unit, students draw initial models at the beginning of the unit (*Lessons 1–4*), revise their models (*Lesson 5*) based on the evidence they collected (*Lessons 2–4*), and make the final consensus models (*Lesson 10*) based on additional evidence (from *Lessons 5–9*). The above strategies ensure the intra-unit coherence of each unit in our learning system, supporting students' science achievement (He, Chen, et al., 2023). To ensure coherence, other members of the team, who did not involve in the primary development work, critiqued the unit for coherence. For instance, they checked to see if students where learning key ideas essential to answer the driving question. We modified the unit sequence based on the evidence from teacher implementation and feedback. For instance, in the development phase, our team employed an online modeling tool (e.g., SageModeler, 2020) in Unit 1 for students to draw and revise models. However, feedback from teachers indicated that the online tool took extra effort for them and their students to learn the drawing tool software. The other 2 units (Units 2 and 4) did not use the online modeling tool; instead, they used hand-drawn models, making the modeling practices across units inconsistent. Thus, our team modified the modeling practices in Unit 1 to hand-drawn modeling but kept the other features (e.g., lesson sequence and driving questions) similar to the initial version.

Our design team used a unit driving question board (Weizman et al., 2008) and an activity summary board (Touitou et al., 2018) to scaffold instructions and support teacher classroom teaching and student learning (see Stage 4c). The open-ended driving question is the core of PBL instruction as it drives student wonderment and ownership in the unit and helps ensure coherence. At the beginning of a unit, the unit driving question and

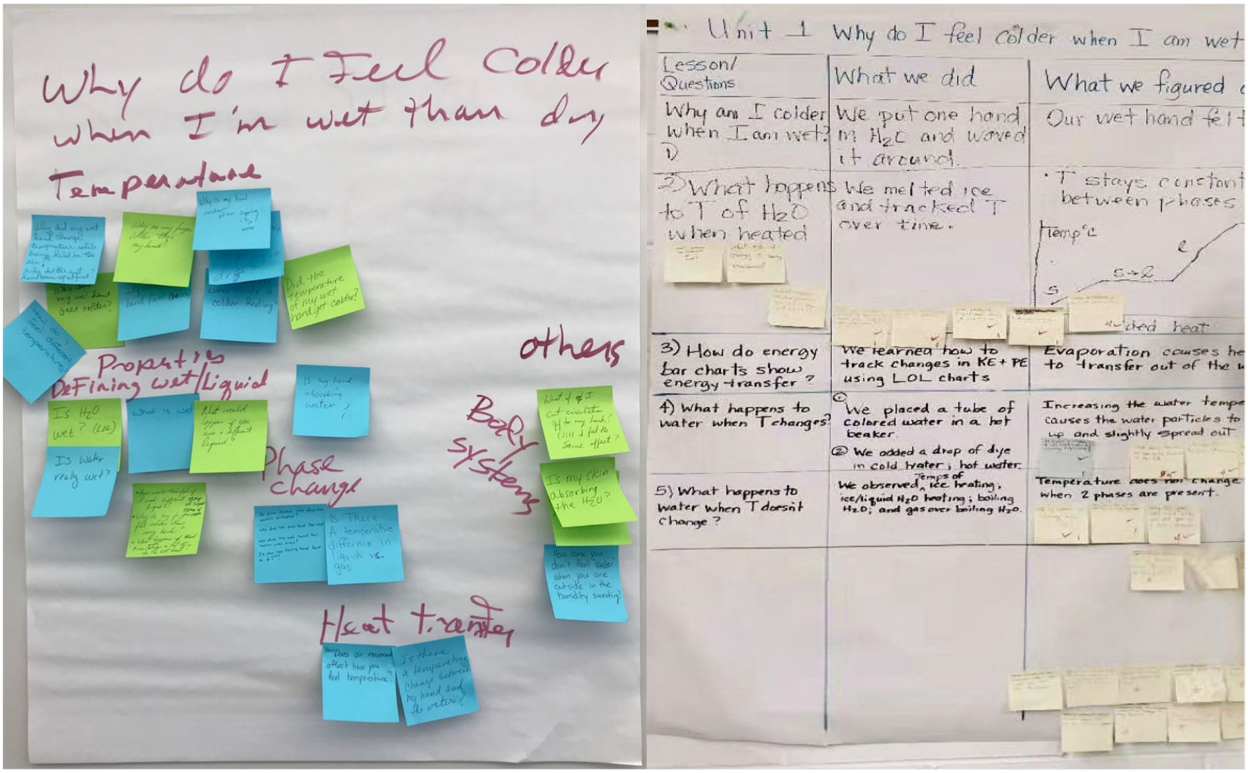


Fig. 4 The Driving Question Board (Left) and Activity Summary Board (Right) in Unit 1

anchoring phenomenon provides students with opportunities to ask sub-questions. Teachers have students ask as many questions as they can. Students generate related sub-questions using sticky notes or on the whiteboard, categorize similar questions, and clarify them. For instance, students generate six categories for grouping their sub-questions related to Unit 1 driving questions: temperature, properties, phase change, heat transfer, body system, and others (see the driving question board on the left side of Fig. 4). Students generated new sub-questions and modified those categories as they gained a deeper understanding of core ideas related to the anchoring phenomenon. The *CESE* design team expects students to figure out the questions they raise as they engage in the learning activities across the lesson sequence of the unit. In the development phase (2015–2017), our team collaborated with teachers and brainstormed all possible sub-questions students might raise in classrooms related to the unit driving question. Those sub-questions were carefully selected and arranged to create lesson-level driving questions. Our pre-developed sub-questions with potential categories could support teachers' instructional practices in the classroom, mainly when students could not generate sub-questions as expected. Moreover, we included students' actual sub-questions from the prior year's classrooms in the updated version of our units in the efficacy and maturation phases.

Along with the driving question board, the activity summary board provides a visual reminder of what students have learned and monitors how their learning progresses across the lessons. The activity summary board consistently checks two questions at the end of each lesson: What have we done? And what did we figure out? (See the activity summary board on the right side of Fig. 4). To wrap up each lesson, the teacher works with students to summarize the disciplinary core ideas, science practices, and crosscutting concepts they experienced to make sense of the lesson-level phenomenon and driving question. Throughout the unit, students move the questions they figured out from the driving question board to the activity summary board. The activity summary board presents students' tangible understanding over time across the unit, which provides evidence to the teacher with respect to student learning and provides evidence for our team to modify learning activities and scaffold instruction. Students' consensus takeaways on the activity summary board provide feedback to the teacher with respect to student learning and support our team in revising the associated evidence statements to the lesson-level learning performance goals.

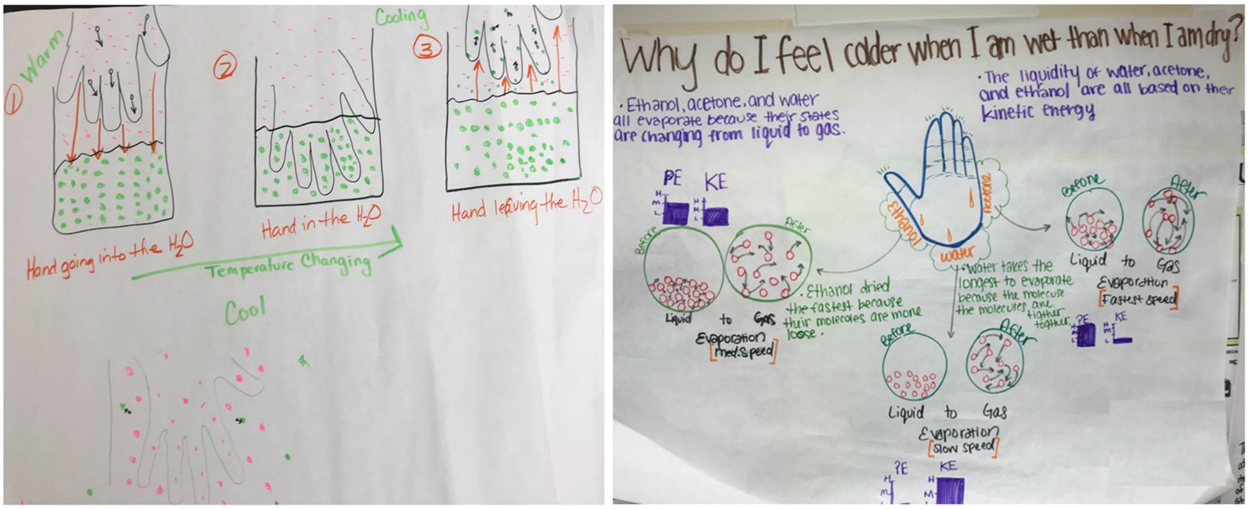


Fig. 5 Student Unit 1 Initial (Left) and Final (Right) Models

Assessment tasks and artifacts

Assessment is another centerpiece of the *CESE* learning system. In Stage 5, our design team intentionally developed two types of assessment tasks, classroom-embedded and post-unit assessments, to improve student learning by diagnosing students' challenges and providing essential information for teachers to make instructional decisions.

Classroom-embedded assessment. Classroom-embedded assessments support student learning in classrooms across a single unit (He, Zhai, et al., 2023). The classroom-embedded task scenarios align with the anchoring phenomenon or lesson-level driving question in each lesson (see Stage 5a). In our system, the tasks include the lesson-level student activity sheets that allow students to record their data, display diagrams, write explanations, and develop and revise models (see Stage 5b). Teachers usually collected student responses on the activity sheets as ongoing artifacts to diagnose students' understanding and challenges in each lesson (see Stage 5c). Figure 5 presents students' initial and final models to explain the Unit 1 driving question. The group of students drew the initial model (the left one in Fig. 5) in Lesson 1 after they experienced the phenomenon of evaporative cooling in their classrooms. The initial model shows the water movement and temperature change when putting and withdrawing a hand in the water. Although the initial model presents the temperature change, it does not explain why hands feel colder when removed from the water. After students experienced several related phenomena in our unit, they collected evidence to develop a deeper understanding of the relationship between energy transferred into or out of the system, and particle motion. They then shared their models with the class and revised them based on feedback and updated knowledge. At the end of Unit 1 (i.e., *Lesson 10*), the students drew a final model to make sense of the unit phenomena (see the final model on the right of Fig. 5). The final model shows the movement of water molecules and their associated energy change in the skin and water during the process. During the efficacy study, we found that the majority of classrooms kept developing and revising the phenomenon (putting one hand in the water and another hand dry) they experienced in *Lesson 1*, which can be seen from the hand in the initial and final models (see Fig. 5). While it is impressive to see students' development across the unit, we modified the *Lesson 10* classroom assignment, making the final model to explain the unit driving question (i.e., *Why do I feel colder when I am wet than when I am dry?)* As such, the unit does not require a specific tangible object in students' final models and explanations.

We also modified some activity sheets to enhance the coherence among lessons within a unit. For instance, in Unit 1, "*Lesson 2: what happens to water when it is heated?*", the lesson goal is to conduct an experiment and analyze and interpret data about temperature changes that occur in water in various phases (i.e., from ice to water, and vapor). The investigation in this lesson is vital for the whole unit as it allows students to engage in first-hand lab experience to collect and analyze data as evidence to support the development and revision of models.

Our design team developed an initial activity (see Table 4 initial version) with a step-by-step procedure (with a thermometer inside the ice) to conduct a lab experiment and asked students to record data in a table

Table 4 The Initial and Final Version of Unit 1 Lesson 2 Activity Sheet (Selected)

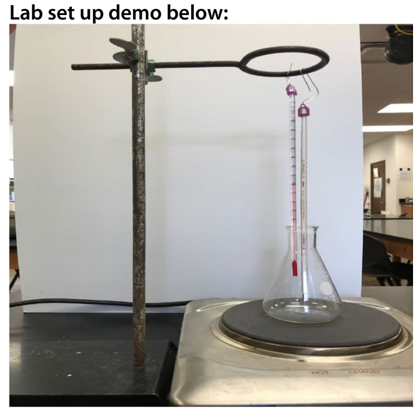
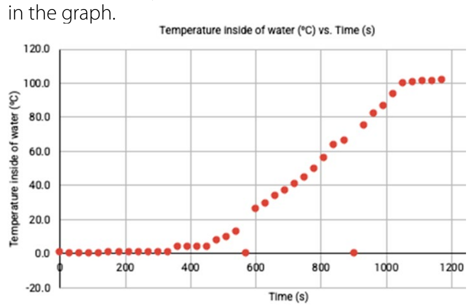
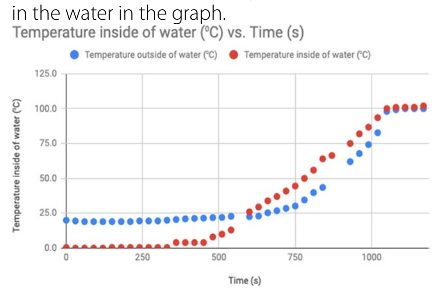
<table>
<tr><th></th><th>Initial version</th><th>Final version</th></tr>
<tr><td>Step-by-step procedure</td><td>Step 1. Put your protective glass and heat-protection gloves.
Step 2. Place about 400 ml of crushed ice in the beaker (it should be about 2/3 full).
Step 3. Insert the thermometer or temperature probe and record the starting temperature.
Step 4. Carefully place the beaker above the Bunsen burner or hot plate.
Step 5. Start recording the time while filling out the data on the table on page 2 of the Lesson 2 activity sheet. Before each data recording, stir the water for 5 seconds using the stirring rod.
Step 6. At the end of the data collection, carefully turn off the Bunsen burner or hot plate. When the beaker cools down, dispose of the water and clean the lab station.</td><td>Step 1. Put on protective glasses and heat-protection gloves.
Step 2. Place about 50 mL of crushed ice in the beaker or flask (about 1/3 full).
Step 3. Place some pieces of boiling stones in the beaker or flask. (if available).
Step 4. Insert one thermometer (or temperature probe) in ice (but do not touch the bottom of the beaker or flask) and another suspended above inside the beaker or flask. Record the starting temperatures both in and out of the ice.
Step 5. Carefully place the beaker above the Bunsen burner or hot plate.
Step 6. Start recording the time while filling out the data in the table below. Gently stir the mixture constantly to ensure you get an accurate system temperature.
Step 7. At the end of the data collection, carefully turn off the Bunsen burner or hot plate. When the beaker cools down, dispose of the water and clean the lab station.
Lab set up demo below:
</td></tr>
<tr><td>Data collection table</td><td>
<table>
<tr><th>Time (sec.)</th><th>Temperature (°C)</th><th>What happens to the water?</th></tr>
<tr><td>0- start</td><td></td><td></td></tr>
<tr><td>30</td><td></td><td></td></tr>
<tr><td>60</td><td></td><td></td></tr>
<tr><td>90</td><td></td><td></td></tr>
<tr><td>120</td><td></td><td></td></tr>
<tr><td>150</td><td></td><td></td></tr>
<tr><td>180</td><td></td><td></td></tr>
<tr><td>210</td><td></td><td></td></tr>
<tr><td>240</td><td></td><td></td></tr>
<tr><td>270</td><td></td><td></td></tr>
<tr><td>300</td><td></td><td></td></tr>
<tr><td>Cont. until boiling...</td><td></td><td></td></tr>
</table>
</td><td>
<table>
<tr><th>Time (sec.)</th><th>Temperature inside of water (°C)</th><th>Temperature outside of water (°C)</th><th>What happens to the water?</th></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td>300</td><td></td><td></td><td></td></tr>
<tr><td>Cont. until boiling...</td><td></td><td></td><td></td></tr>
</table>
</td></tr>
</table>

Table 4 (continued)

	Initial version	Final version
Data Graph	Now, draw a graph of Temperature vs. Time on the data you have collected in the box below. Also, point out the changes you have observed in the water in the graph. 	Now, draw a graph of Temperature vs. Time on the data you have collected in the box below. Also, point out the changes you have observed in the water in the graph. 

and draw a temperature-time graph to show the temperature changes over time. Our efficacy study found that students could explain the temperature trends from ice to water to gas. However, they failed to explain the role of energy when water turns into a gas (in *Lesson 5*) because the lab experiment did not provide evidence with data to show that the temperature of boiling water and vapor are the same. Moreover, our participating teachers had experienced challenges in planning and carrying out this lab experiment, resulting in unpredictable data trends. Therefore, we modified the activity (see Table 4 final version) by 1) providing a lab set-up demonstration, 2) adding another thermometer suspended above the ice, and 3) adding another column in the data table to record temperature changes outside of water. We found that students could draw a graph with two lines to show the difference in the temperature changes between the ice (water) and the gas.

Post-unit assessment. A post-unit assessment evaluates student performance at the end of a unit. Tasks in a post-unit assessment should share the same learning performance goals as the unit but are situated in proximal scenarios to assess learners' ability to transfer their knowledge and skills (Ruiz-Primo et al., 2002). In our system, the design team developed the post-unit assessment tasks and associated rubrics aligned with the same NGSS PEs for each unit by modifying the assessment design process (Harris et al., 2019; He, Zhai, et al., 2023). The design team followed the three steps in Stage 5 (see Fig. 2) to design assessment tasks. Using the unpacking materials of three dimensions, the design team articulated task learning performances and associated evidence statements. Next, the design team determined a task design blueprint (see Stage 5b) consisting of characteristic and variable task features incorporating equity and fairness considerations (Rose et al., 2005; Rose & Meyer, 2006). To align with the final artifacts of the units, the design team deliberately employed "*Developing and Using Models*" and "*Cause and Effect*" as the focal practice and the crosscutting concept in all task learning performances (e.g., He, Chen, et al., 2023; Li et al., 2021; Li, Chen, et al., 2023). To achieve the tasks, students demonstrate their proficiencies by drawing a microscope-level model and using it to explain the anchoring phenomenon. Table 5 shows the task scenarios and associated learning performances in the four post-unit assessments (see Stage 5a). Our prior study reported the three post-unit assessments (Units 1, 2, and 4) (He, Chen, et al., 2023). Our results show how the *CESE* team monitored student development across the 3 units by cumulative predictions of their science achievement on a third-party end-of-year summative test.

This paper takes one of the tasks in Unit 1 post-unit assessment as an example (see Table 6) to elaborate on our development and revision process in Stage 5 (See Fig. 2). In the development phase, our design team developed a task scenario of placing ice cubes on a hot summer day and provided a temperature-time graph and a data table of observation records. Prompt 2 asks students to draw a molecular-level model to show the water molecules at point B (i.e., the ice is starting to melt). The prompt in the task was designed based on the learning performance and associated evidence statement (see Table 7). To scaffold students' responses, our team provided a model of water molecules at point A (i.e., ice; see the initial version in Table 6). Accordingly, students' drawn responses are expected to show the vibration of water molecules but remain in the same positions as point A. Our design team developed a holistic rubric of three levels (proficient, developing, and beginning) to score students' drawn responses (see Stage 5c). However, we found a significant issue from

Table 5 The Overview of the Four Post-Unit Assessment Tasks

	Task Scenarios	3D Task Learning Performances
Unit 1: Evaporative Cooling	*Scenario 1:* Placing ice cubes on a hot summer day (Tasks Q1-Q4)	Students develop and use a molecular-level model to explain that ice cubes' temperature increases but remain solid when they absorb energy from their surroundings. (Q1 and Q2) Students develop and use a molecular-level model to explain the phase change of ice cubes (solid to liquid), but the temperature remains the same when they absorb energy from their surroundings. (Q3 and Q4)
	Scenario 2: Observing ice cubes and rubbing alcohol at −6 degrees Celsius (Tasks Q5-Q6)	Students develop and use molecular-level models to explain why the states of ice cubes and rubbing alcohol are not the same when they are placed at the same temperature (−6 degrees Celsius). (Q5 and Q6)
Unit 2: Periodical Table	*Scenario 1:* Alkali metals reacting in water (Tasks Q1-Q3)	Students develop and use an atomic-level model to explain and predict the relative ease with that alkali metals (sodium, lithium, and rubidium) react with water. (Q1, Q2, and Q3)
	Scenario 2: Magnesium reacting in water (Tasks Q4-Q6)	Students develop and use an atomic-level model to explain why magnesium is less reactive than sodium when placed in water. (Q4, Q5, and Q6)
Unit 3: Combustion Reactions	*Scenario 1: Burning methane in a beaker (Task Q1-Q4)*	Students develop and use an atomic-level model to explain the changes in substances and energy before and after a simple chemical reaction (Q1 and Q2) Students use a diagram model to explain why a combustion reaction releases energy to make things hot. (Q3) Students use a diagram model to explain why we need to add energy to start a combustion reaction (Q4)
Unit 4: Conservation of Mass and Atom	*Scenario 1:* Burning iron wool in the air (Tasks Q1-Q4)	Students develop and use an atomic-level model to explain how mass and atoms are conserved before and after a simple chemical reaction. (Q2 and Q3) Students construct and use a mathematical representation model to explain why the mass of the product obtained from burning iron wool is larger than the iron wool before the burning. (Q1 and Q4)

analyzing a thousand students' drawn responses: many students either copied the same model as point A or presented a molecular model combining the organized and disorganized structure of water molecules. Our team realized this issue is because of the inappropriate presentation of point B in the graph and data table (A-B: Mostly solid ice). In the final version (see Table 6 in the right column), we modified the graph to move point B at −2 degrees Celsius, and the observation of A to B in the table is solid ice. Instead of scaffolding with an exemplar model, we provided an image of actual ice so that students could use what they had learned in the unit to draw a molecular-level model (see an example in Table 7, in the right column). While a holistic rubric only provides evaluation information, we created an analytic rubric, including components, relationships, and casual explanations for scoring students' drawn responses (He, Chen, et al., 2023). Accordingly, the entire score on the post-unit assessment can evaluate students' overall performance after completing a unit. The scores on the tasks in the post-unit assessment could provide teachers with detailed diagnostic information on students' challenges to adjust their teaching plans in the following units.

Professional learning

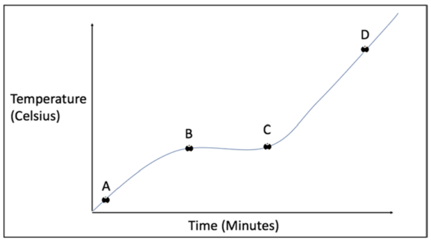
Our team provided professional learning to support participating teachers in implementing the unit and assessment materials in our *CESE* learning system (Stage 6). Our professional learning occurred in three different situations: a) in-person workshops before implementing the units, b) virtual check-in meetings during the units, and c) asynchronous communication via emails, phone calls, Slack, and other chat programs. We provided separate in-person workshops before the implementation of each unit. The first three-day PL session always occurred in the summer. Day 1 introduced the overview of our project, the logistics of data collection, the vision of NGSS,

Table 6 The Initial and Final Version of Assessment Tasks in Unit 1 Post-Unit Assessment

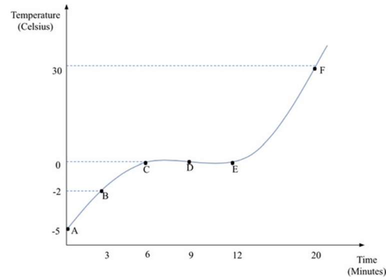
Initial version	Final version

Task scenario: Placing ice cubes on a hot summer day

To investigate how the temperature of ice changes on a hot summer day, Kellie put three thermometers in an ice cube tray, filled the tray with water, and put the tray in the refrigerator overnight. In the morning, she took the tray outside and recorded the temperature of three ice cubes and her observations of what happened to each cube. The average temperature change of the three cubes is reported in the graph below.



Kellie wrote her observations in the table.

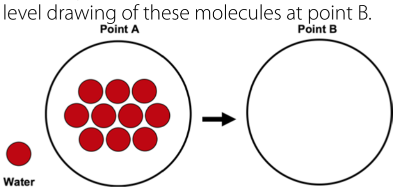


Kellie wrote her observations in the table.

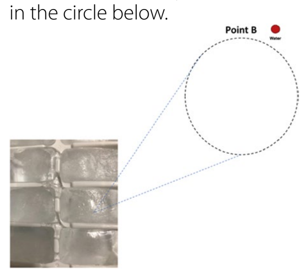
Time Point	Observation	Time Point	Observation
A - B	Mostly solid ice	From A to B	Solid ice
B - C	The mixture of ice and water	From B to D	From solid ice to a mixture of solid ice and liquid water
C - D	Mostly liquid water	From D to E	From the mixture of solid ice and liquid water to all liquid water
		From E to F	All liquid water

Kellie was curious why the water sometimes increased in temperature, while at other times, it did not change its temperature but changed from solid to liquid. Answer the following questions.

Question 2. Below is a molecular-level drawing of water molecules at point A. Using your explanation from the previous question, construct a molecular-level drawing of these molecules at point B.



Question 2. If you have a very powerful microscope and you could see what the ice cube is made up of at point B, and how is it structured? What would you see? Please draw a microscope-level model in the circle below.

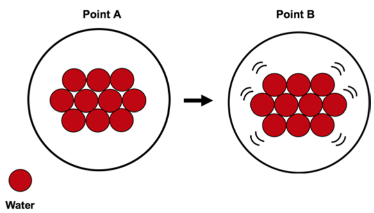
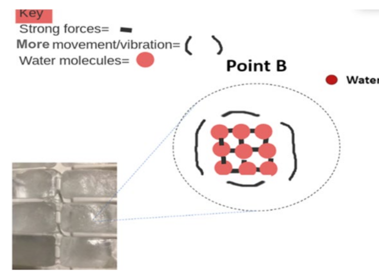


and the *Framework of K-12 Science Education* (NRC, 2012). Day 2 overviewed the critical features of the PBL curriculum materials and the entire learning system. Day 3 introduced a specific unit and allowed teachers to experience the main activities in the unit as students may have in the classrooms.

Table 8 presents the agenda for Unit 1 professional learning. Using the PBL approach, the *CESE* team created a professional learning session driving question: "*How can we create Engaging Learning Environments that support students in making sense of phenomena through PBL?*" The *CESE* professional development team specified four learning goals for this PL session and developed professional learning activities to address these learning goals. In Unit 1 professional learning session, teachers first experienced the anchoring phenomenon (i.e., the cooling effect of water evaporation on one's hands), asked questions related to the phenomenon, and posted their questions on a driving question board. Next, they engaged in data collection and analysis using temperature-time diagrams. Finally, they drew and shared the initial models and revised them based on peers' feedback. Such unit-focused PL activities allowed teachers to gain experience teaching the *CESE* units in classrooms.

To ensure the success of implementing the units, we gradually modified our professional learning plans over

Table 7 The Initial and Final Version of Scoring Rubrics for the Assessment Tasks in Unit 1 Post-Unit Assessment

	Initial version	Final version
Learning performance	Students develop a molecular-level model to explain that ice cubes' temperature increases but remain solid when they absorb energy from their surroundings. (Q2)	
Evidence statement	Student's model shows the water molecules at point B were connected and organized as a solid at a macroscope level, but they vibrated more than before.	
Exemplar response		
Criteria	Holistic rubric Proficient (3): Student depicts water molecules in roughly the same organization vibrating more quickly. Developing (2): Student re-draws the original drawing without indicating vibration. Beginning (1): Student draws a disorganized mass or particles that would be characteristic of a particulate-level liquid or gas representation	Analytic rubric Components (C): The model includes the identification and specification of appropriate and essential components, including both visible and invisible. • C1-Water molecules • C2-Inter-molecule interactions • C3-Vibration of molecules Relationships (R): The model includes representations or descriptions indicating how various components within the model are related. • R1: Organized structure with all components (include C1, C2, and C3 with clearly labels) • R2: Water molecules are connected at a close distance Causal Explanations (CE): The model is used to explain or predict phenomena. • CE1: Model Includes all components (C1, C2 and C3) and relationships (R1 and R2). • CE2: Keys or short written reasoning of Components and Relationships. (e.g., the model shows water molecules in the ice vibrating more and connected in an organized structure.)

the years. We increased the unit-specific sessions whenever we developed new units. We added a session on Day 1 that invited a panel of returning teachers to share their teaching experiences. In addition, we invited our lead teacher to co-facilitate the unit-specific session. They usually led lab investigation and group discussion and shared their instructional strategies with new participating teachers. Moreover, we were excited to see our participating teachers form a professional learning community by creating a shared folder in Google Drive and sharing resources in the professional learning sessions and during their implementation of our units. Due to the outbreak of COVID-19, we could not provide in-person sessions as before. Our team moved all sessions virtually and increased bi-weekly check-in meetings to support teachers. During the virtual sessions, we used online tools such as *Jamboard* and *Zoom* breakout rooms, which could effectively facilitate teachers' idea sharing and group discussion.

Conclusion and future direction

Promoting student knowledge-in-use has significantly gained attention in worldwide K-12 science education. Although science standards in many countries (e.g., Finnish National Board of Education, 2016; MoE, 2017; NGSS Lead States, 2013) have addressed this emerging trend, challenges remain for developing a standards-aligned and coherent learning system to support student knowledge-in-use development effectively and sustainably. Moreover, the lack of rigorous research design and empirical evidence impede stakeholders, researchers, and practitioners from enacting such standards-aligned curriculum and teaching materials in classrooms.

The *CESE* project addresses the above challenges and has put effort into developing, implementing, and testing a learning system of high school chemistry and physics, consisting of teacher and student curriculum and instructional materials, assessment, and professional learning. The *CESE* learning system was designed based on the

Table 8 Unit 1 Professional Learning Outline

PL driving question: How can we create Engaging Learning Environments that support students in making sense of phenomena through PBL?
Session Learning Goals:
1. Develop strategies to support students in 3D learning using MSS(NGSS) and PBL.
2. Form a community of learners to enact PBL in the classroom to support students in deep sensemaking.
3. Explore technology resources that support modeling to be used by teachers and students in *ESBS* units.
4. Explore the role of the teacher in supporting discourse in *ESBS* units.

Time	Activity	Goal addressed
8:00–8:30	Welcome/ Introductions	2
8:30–9:00	Opening Discussion: How has the New MSS changed your teaching?	1
9:00–9:30	Project Overview	1, 2
9:30–10:00	• Experiencing PBL: *Evaporative Cooling* Unit • Experiencing Phenomena and Introducing the Driving Question • Experiencing hands in water bowls- developing the DQ board (include the question about different liquids)	1, 2
10:00–10:45	Using Scientific Practices: • Experiencing evaporation of two liquids ○ Performing the liquid drop activity in pairs • Developing models to explain the difference between the evaporation rate of liquids ○ Drawing models in pairs ○ Sharing models with another pair ○ Sharing several models with the whole group	1, 2
10:45–11:00	BREAK	
11:00–11:45	Debriefing: Features of Project-based Learning	1, 2
11:45–12:15	How does PBL support student learning: PBL and 3-Dimensional learning?	1, 2
12:15–1:00	WORKING LUNCH	
1:00–1:30	Session Driving Question: How can we create Learning Environments that support students in making sense of phenomena through PBL?	3, 4
1:30–2:30	Developing Optimal Learning Environments through Discourse Teacher/Student Talk Moves Turn and talk- Practicing talk moves: How might your classroom environment be different with PBL?	3, 4
2:30–2:45	BREAK	
2:45–3:15	Teacher Panel	1, 2
3:15–3:30	Wrap-Up- Revisiting our Driving Question- New Questions	1, 2

theoretical perspectives of 3D learning (NRC, 2012), PBL (Krajcik & Czerniak, 2018), situated engagement (Schneider et al., 2016; Schmidt et al., 2018), and learning progression to ensure coherence (Fortus & Krajcik, 2012). The *CESE* learning system is an effective intervention for supporting students' knowledge-in-use development based on the results from our efficacy study (He, Chen, et al., 2023; Schneider et al., 2022). The evidence shows that the *CESE* learning system is theory-driven and empirically validated. The *CESE* materials have transformed standards into classroom materials to promote students' knowledge-in-use in high school chemistry.

In this paper, we illustrated our design framework and elaborated on the process for iteratively designing, testing, and revising the *CESE* learning system. Our paper presents a holistic landscape that shows how to transform national science standards (in our case, the NGSS) into curriculum materials usable by classroom teachers and students. The *CESE* principled design approach to developing and testing standards-aligned learning systems would contribute to the global science education community as follows. First, curriculum developers would benefit from our design principles and framework for designing standards-aligned learning systems in their educational contexts. Accordingly, they may need to localize our design process based on their needs. However, we encourage their localization of our design process to keep our essential components, including standards unpacking, curriculum, assessment, and professional learning. The specific sub-components (e.g., 2a, 2b, and 2c in Stage 2, see Fig. 2) can be modified or adjusted to meet the features of learning performance goals in their standards. Second, our project research design (He, Chen, et al., 2023; Schneider et al., 2022) would be beneficial for researchers who are interested in creating and testing the intervention of a research-based learning system to promote student development of complex learning outcomes (e.g., knowledge-in-use). Third, this paper provides an exemplary model of creating a collaborative learning community with researchers, teachers, state leaders, and administrators to support students' long-term development to meet the standards-aligned learning goals. Such theory-driven and empirically validated learning systems would support teachers in transforming their local standards into curriculum materials.

The *CESE* learning system would potentially impact local policies and classroom practices in K-12 science education. Our efficacy study provides robust evidence using a cluster randomized trial experimental design approach and shows that our *CESE* learning system could impact students' end-of-year science achievement in Michigan and California (Schneider et al., 2022). Our future direction is to scale up the intervention of our learning system and generalize the main effects of our intervention to a broader population across the nation. Our learning system has not yet covered all NGSS PEs in high school physical sciences. As such, we aim to continually develop more PBL units into the existing learning system to establish a full-covered standards-aligned learning system to support students' knowledge-in-use in high school physical sciences. Developing long-term 3D learning progressions can guide intra- and inter-unit coherence of standards-aligned learning systems (He et al., in press, a, b). As a centerpiece of standards-aligned learning systems, thorough design and implementation of classroom-embedded assessments alongside a 3D learning progression are needed to collect students' ongoing artifacts for tracking their progress toward the targeted learning goals in science standards. A rigorous classroom-embedded post-unit and end-of-year assessment system would diagnose students' challenges and evaluate their progress within and between units in a learning system. In addition, we recommend that future studies consider using innovative technologies (e.g., artificial intelligence and data mining) to assist and enhance such learning systems (He et al., in press, b). AI-based technologies (e.g., large language models and machine learning) can capture students' challenges in their responses to classroom-embedded assessment tasks (e.g., Li, Liu, & Krajcik, 2023). Accordingly, teachers and students would receive immediate and meaningful feedback to adjust their teaching and learning strategies to achieve the targeted learning goals (He et al., in press, c). While recognizing the ethics and transparency issues in applying AI technologies in education (Li, Reigh, et al., 2023), we envision generative AI as a promising and unexplored area for enhancing standards-aligned and coherent learning systems to better enhance student knowledge-in-use development across time.

Abbreviations

NGSS	Next Generation Science Standards
NRC	National Research Council
3D learning	three-dimensional learning
PBL	project-based learning
PE	performance expectation
CESE	Crafting Engagement in Science Environments

Acknowledgements
The authors thank the initial curriculum and assessment developers (Tom Bielik, Israel Touitou, Ryan Stowe), lead teachers (Sandra Erwin, Will Paddock), and all participating teachers. The materials are based upon work supported by National Science Foundation (OISE- 1545684; PIs Barbara Schneider & Joe Krajcik). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Authors' contributions
PH drafted, revised, and finalized the manuscript. JK commented and edited the manuscript. BS commented and improved the manuscript. All authors read and approved the final manuscript.

Author information
1. Peng He is an Assistant Professor at CREATE for STEM Institute at Michigan State University, USA.
2. Joe Krajcik is the co-Principal Investigator on the Crafting Engaging Science Environment Project, the Lappan-Philips Professor of Science Education, University Distinguished Professor, and Director of CREATE for STEM Institute at Michigan State University, USA.
3. Barbara Schneider is the Principal Investigator on the Crafting Engaging Science Environment Project, the John A. Hannah University Distinguished Professor at Michigan State University College of Education, USA.

Funding
National Science Foundation (OISE- 1545684).

Availability of data and materials
All materials are Creative Commons Open Sources and are freely available for public use. The materials have the designation of Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0). Interested individuals can access the materials at https://sites.google.com/a/msu.edu/craftingengagingscience/. Data sharing is not applicable to this article as no datasets were generated or analyzed in this study.

Declarations

Competing interests
The authors declare that they have no competing interests.

Received: 24 January 2023 Accepted: 22 November 2023
Published online: 20 December 2023

References

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences*, *13*(1), 1–14.
Bielik, T., Finnie, K., Peek-Brown, D., Klager, C., Touitou, I., Schneider, B., & Krajcik, J. (2022). High school teachers' perspectives on shifting towards teaching NGSS-aligned project based learning curricular units. *Journal of Science Teacher Education*, *33*(4), 413–434.
Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist*, *26*(3–4), 369–398.
Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn (Vol. 11)*. National Academies Press.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*(1), 32–42.
Csikszentmihalyi, M., & Schneider, B. (2000). Becoming adult: How teenagers prepare for the world of work. Basic Books.
Condliffe, B. (2017). *Project-based learning: A literature review*. Working Paper. MDRC.
Duncan, R. G., Krajcik, J. S., & Rivet, A. E. (Eds.) (2017). *Disciplinary core ideas: Reshaping teaching and learning*. National Science Teachers Association: NTSA Press.
Duschl, R., Maeng, S., & Sezen, A. (2011). Learning progressions and teaching sequences: A review and analysis. *Studies in Science Education*, *47*(2), 123–182.
Finnish National Board of Education (2016). *National core curriculum for basic education 2014*. Helsinki: National Board of Education.
Fortus, D., & Krajcik, J. (2012). Curriculum coherence and learning progressions. In *Second international handbook of science education (pp. 783–798)*. Springer.
Fulmer, G. W., Tanas, J., & Weiss, K. A. (2018). The challenges of alignment for the next generation science standards. *Journal of Research in Science Teaching*, *55*(7), 1076–1100.
Geier, R., Blumenfeld, P. C., Marx, R. W., Krajcik, J. S., Fishman, B., Soloway, E., & Clay-Chambers, J. (2008). Standardized test outcomes for students engaged in inquiry-based science curricula in the context of urban reform. *Journal of Research in Science Teaching*, *45*(8), 922–939.
Haas, A., Januszyk, R., Grapin, S. E., Goggins, M., Llosa, L., & Lee, O. (2021). Developing instructional materials aligned to the next generation science standards for all students, including English learners. *Journal of Science Teacher Education*, *32*(7), 735–756.
Harris, C. J., Krajcik, J. S., Pellegrino, J. W., & DeBarger, A. H. (2019). Designing knowledge-in-use assessments to promote deeper learning. *Educational Measurement: Issues and Practice*, *38*(2), 53–67.
He, P., Chen, I.-C., Touitou, I., Bartz, K., Schneider, B., & Krajcik, J. (2023). Predicting student science achievement using post-unit assessment performances in a coherent high school chemistry project-based learning system. *Journal of Research in Science Teaching*, *60*(4), 724–760.
He, P. Shin, N. Kaldaras L., & Krajcik, J. (in press, B). Integrating artificial intelligence into learning progression-based learning systems to support student knowledge-in-use: Opportunities and challenges. In Jin, H., Yan, D., & Krajcik, J. *Handbook of research in science learning progressions.*
He, P., Shin, N., & Krajcik, J. (in press, A). Developing three-dimensional learning progressions of energy, interaction, and matter at middle school level: A design-based research. In Jin, H., Yan, D., & Krajcik, J. *Handbook of research in science learning progressions.*
He, P., Shin, N., Zhai, X., & Krajcik, J. (in press, c). A design framework for integrating artificial intelligence to support teachers' timely use of knowledge-in-use assessments. In Zhai, X & Krajcik, J. Uses of artificial intelligence in STEM education. Oxford University Press.
He, P., Zhai, X., Shin, N., & Krajcik, J. (2023). Applying Rasch measurement to assess knowledge-in-use in science education. In X. Liu, & W. J. Boone (Eds.), *Advances in applications of Rasch measurement in science education. Contemporary trends and issues in science education*, (vol. 57). Cham: Springer. https://doi.org/10.1007/978-3-031-28776-3_13.
He, P., Zheng, C., & Li, T. (2022). Development and validation of an instrument for measuring Chinese chemistry teachers' perceived self-efficacy towards chemistry core competencies. *International Journal of Science and Mathematics Education*, *20*(7), 1337–1359.
Hinojosa, T., Rapaport, A., Jaciw, A., LiCalsi, C., & Zacamy, J. (2016). *Exploring the foundations of the future STEM workforce: K–12 indicators of postsecondary STEM success*. U.S. Department of Education.
Inkinen, J., Klager, C., Juuti, K., Schneider, B., Salmela-Aro, K., Krajcik, J., & Lavonen, J. (2020). High school students' situational engagement associated with scientific practices in designed science learning situations. *Science Education*, *104*(4), 667–692.
Inkinen, J., Klager, C., Schneider, B., Juuti, K., Krajcik, J., Lavonen, J., & Salmela-Aro, K. (2019). Science classroom activities and student situational engagement. *International Journal of Science Education*, *41*(3), 316–329.
Jin, H., Mikeska, J. N., Hokayem, H., & Mavronikolas, E. (2019). Toward coherence in curriculum, instruction, and assessment: A review of learning progression literature. *Science Education*, *103*(5), 1206–1234.
Krajcik, J., McNeill, K. L., & Reiser, B. J. (2008). Learning-goals-driven design model: Developing curriculum materials that align with national standards and incorporate project-based pedagogy. *Science Education*, *92*(1), 1–32.
Krajcik, J., Schneider, B., Miller, E., Chen, I.-C., Bradford, L., Baker, Q., … Peek-Brown, D. (2023). Assessing the effect of project-based learning on science learning in elementary schools. *American Education Research Journal*, *60*(1), 70–102.
Krajcik, J. S., & Czerniak, C. M. (2018). *Teaching science in elementary and middle school: A project-based learning approach*. Routledge.
Li, T., Chen, I., Miller, E., Miller, C., Schneider, B., & Krajcik, J. (2023). The relationships between elementary students' knowledge-in-use performance and their science achievement. *Journal of Research in Science Teaching*, 1–61.
Li, T., Liu, F., & Krajcik, J. (2023) Automatically assess elementary students' hand-drawn scientific models using deep learning of artificial intelligence. Proceedings of the annual meeting of the International Society of the Learning Sciences (ISLS).
Li, T., Miller, E., Chen, I. C., Bartz, K., Codere, S., & Krajcik, J. (2021). The relationship between teacher's support of literacy development and elementary students' modelling proficiency in project-based learning science classrooms. *Education 3-13: International Journal of Primary, Elementary and Early Years Education*, *49*(3), 302–316.
Li, T., Miller, E. A., & Krajcik, J. S. (2023). Theory into practice: Supporting knowledge- in-use through project-based learning. In G. Bansal, & U. Ramnarain (Eds.), *Fostering science teaching and learning for the fourth industrial revolution and beyond (pp. 1–35)*. IGI Global. https://doi.org/10.4018/978-1-6684-6932-3.ch001.
Li, T., Reigh, E., He, P., & Adah Miller, E. (2023). Can we and should we use artificial intelligence for formative assessment in science? *Journal of Research in Science Teaching*, *60*(6), 1385–1389.
Miller, E. C., & Krajcik, J. S. (2019). Promoting deep learning through project-based learning: A design problem. *Disciplinary and Interdisciplinary Science Education Research*, *1*(1), 1–10.
Ministry of Education, P. R. China (2017). *Chemistry curriculum standards for senior high school [普通高中化学课程标准]*. People's Education Press.
Mislevy, R., & Haertel, G. (2006). Implications of evidence-centered design for educational testing. *Educational Measurement: Issues and Practice*, *25*(4), 6–20.
National Center for Education Statistics. (2017). The condition of education 2017. Author. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2017144
National Research Council (2000). *How people learn: Brain, mind, experience, and school: Expanded Edition*. Washington, DC: The National Academies Press.
National Research Council (2006). *Systems for state science assessment*. National Academies Press.
National Research Council. (2007). Taking science to school: learning and teaching science in grades K-8. Washington, DC: The National Academies Press. https://doi.org/10.17226/11625.
National Research Council (2011). *Assessing 21st century skills: Summary of a workshop*. National Academies Press.
National Research Council (2012). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press.
NGSS Lead States (2013). *Next generation science standards: For states, by states*. National Academies Press.
Nordine, J., & Lee, O. (2021). *Crosscutting concepts: Strengthening science and engineering learning*. NSTA Press.
Organization for Economic Cooperation and Development (2019). *PISA 2018 assessment and analytical framework*. OECD Publishing.
Pellegrino & Hilton (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. National Academies Press.
People's Republic of China Ministry of Education. (2014). Opinions on deepening curriculum reform and implementing the fundamental tasks of Lide-Shuren. http://www.moe.gov.cn/srcsite/A26/jcj_kcjcgh/201404/t20140408_167226.html.
Reiser, B. J. (2014). *Designing coherent storylines aligned with NGSS for the K-12 classroom*. Boston, MA: National Science Education Leadership Association Meeting.
Reiser, B. J., Novak, M., McGill, T. A., & Penuel, W. R. (2021). Storyline units: An instructional model to support coherence from the students' perspective. *Journal of Science Teacher Education*, *32*(7), 805–829.
Rose, D. H., & Meyer, A. (2006). *A practical reader in universal design for learning*. Harvard Education Press.

Rose, D. H., Meyer, A., & Hitchcock, C. (Eds.) (2005). *The universally designed classroom: Accessible curriculum and digital technologies*. Harvard Education Press.
Roseman, J. E., Stern, L., & Koppal, M. (2010). A method for analyzing the coherence of high school biology textbooks. *Journal of Research in Science Teaching*, *47*(1), 47–70.
Ruiz-Primo, M. A., Shavelson, R. J., Hamilton, L., & Klein, S. (2002). On the evaluation of systemic science education reform: Searching for instructional sensitivity. *Journal of Research in Science Teaching*, *39*(5), 369–393.
SageModeler [Computer software] (2020). *Concord, MA: The Concord consortium and the CREATE for STEM*. Institute at Michigan State University.
Schmidt, W. H., Wang, H. C., & McKnight, C. C. (2005). Curriculum coherence: An examination of US mathematics and science content standards from an international perspective. *Journal of Curriculum Studies*, *37*(5), 525–559.
Schmidt, J. A., Rosenberg, J. M., & Beymer, P. N. (2018). A person-in-context approach to student engagement in science: Examining learning activities and choice. *Journal of Research in Science Teaching*, *55*(1), 19–43.
Schneider, B., Krajcik, J., Lavonen, J., & Salmela-Aro, K. (2020). *Learning science: The value of crafting engagement in science environments*. Yale University Press.
Schneider, B., Krajcik, J., Lavonen, J., Salmela-Aro, K., Broda, M., Spicer, J., … Viljaranta, J. (2016). Investigating optimal learning moments in US and Finnish science classes. *Journal of Research in Science Teaching*, *53*(3), 400–421.
Schneider, B., Krajcik, J., Lavonen, J., Salmela-Aro, K., Klager, C., Bradford, L., … Bartz, K. (2022). Improving science achievement—Is it possible? Evaluating the efficacy of a high school chemistry and physics project-based learning intervention. *Educational Researcher*, *51*(2), 109–121.
Schwarz, C. V., Passmore, C., & Reiser, B. J. (2017). *Helping students make sense of the world using next generation science and engineering practices*. NSTA Press.
Shin, N., Stevens, S. Y., & Krajcik, J. (2010). Tracking student learning over time using construct-centred design. In S. Rodrigues (Ed.), *Using analytical frameworks for classroom research (pp. 56–76)*. Routledge.
Touitou, I., Barry, S., Bielik, T., Schneider, B., & Krajcik, J. (2018). The activity summary board. *The Science Teacher*, *85*(3), 30–35.
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
Weizman, A., Shwartz, Y., & Fortus, D. (2008). The driving question board. *The Science Teacher*, *75*(8), 33.
Zhao, Y., & Wang, L. (2022). A case study of student development across project-based learning units in middle school chemistry. *Disciplinary and Interdisciplinary Science Education Research*, *4*(1), 1–20.

Publisher's Note