



Augmenting the effect of virtual labs with "teacher demonstration" and "student critique" instructional designs to scaffold the development of scientific literacy

Chen-Chung Liu¹ · Cai-Ting Wen² · Hsin-Yi Chang³  · Ming-Hua Chang¹ · Po-Han Lai² · Shih-Hsun Fan Chiang¹ · Chih-Wei Yang⁴ · Fu-Kwun Hwang⁵

Received: 23 November 2020 / Accepted: 22 November 2021 / Published online: 4 January 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

Virtual labs provide space for students to iteratively test, observe, and revise their understanding so as to improve their scientific literacy. However, one of the challenges that students face is that they need to think and act like scientists so as to be sensitively alert to methodological flaws and various sources of error. This study thus compared the effect of two instructional approaches using a virtual lab to enhance students' scientific literacy. Before students were given the opportunity to conduct science inquiries with the virtual lab, they were required to critique problematic inquiry cases (the critique group) or watch teachers' demonstrations (the teacher demonstration group) before taking part in the inquiry. By analyzing data from 50 middle school students, this study found that the effect of applying virtual labs can be augmented by an instructional design that engages students in critiquing experiments prior to their inquiry with the virtual lab. This study also found a limitation of the use of virtual labs in helping students transfer what they have learned from the teacher's demonstration to new inquiry contexts. A close relation among scientific literacy post-test scores, critiquing performance, and inquiry performance in the inquiry activity was detected, suggesting that student critiquing prior to inquiry is in alignment with the goal of developing students' inquiry skills and scientific literacy with virtual labs.

Keywords Virtual lab · Scientific literacy · Critique · Inquiry

✉ Hsin-Yi Chang
hychang@ntnu.edu.tw

¹ Department of Computer Science and Information Engineering, National Central University, Taoyuan City, Taiwan

² Graduate Institute of Network Learning Technology, National Central University, Taoyuan City, Taiwan

³ Program of Learning Sciences and Institute for Research Excellence in Learning Sciences, National Taiwan Normal University, #162, Sec.1, Heping E. Rd., Taipei City 106, Taiwan

⁴ Graduate Institute of Educational Information and Measurement, National Taichung University of Education, Taichung, Taiwan

⁵ Department of Physics, National Taiwan Normal University, Taipei City, Taiwan

Introduction

Educators and researchers have been advocating the critical role of fostering students' scientific literacy in contextualized environments, such as engaging students in project-based learning in the area of STEAM or STEM education (Adriyawati et al., 2020; Kusumastuti et al., 2019), or in argument interventions (Cavagnetto, 2010). Such an emphasis on scientific literacy not only aims to help students learn predefined knowledge, but also to develop students' competencies in constructing scientific knowledge with scientific methods (OECD, 2016; Osborne, 2014). The OECD PISA Framework defines scientific literacy as the ability to explain phenomena scientifically, evaluate and design scientific inquiry, describe and appraise scientific investigations and propose ways of addressing questions scientifically. However, direct teaching of science knowledge is still frequently applied in school systems all around the world. In science classrooms where direct teaching is the main pedagogy, students tend to follow teachers' instruction to understand the pre-defined science concepts without the opportunity to experience the science practice and thus develop their scientific literacy. Researchers have advocated that learning in decontextualized didactic environments may impede the transfer of the learning experience to other situations, and literacies should be cultivated in authentic scientific practice (Charney et al., 2007).

Science laboratories in this sense play an important role in providing students with access to the science practice to promote students' literacy in scientific practical skills and understanding of how science and scientists work (Hofstein & Mamlok-Naaman, 2007). However, due to the orchestration constraints such as time constraints and availability of instruments (Vergara et al., 2017), physical laboratories are not widely available or frequently adopted in many schools. Virtual labs that virtually represent and demonstrate scientific phenomena with computers become extraordinarily important to developing scientific literacy particularly during the disruption caused by COVID-19. However, simply giving students virtual labs may not lead to favorable effects due to the complexity of scientific inquiry and virtual labs (Akaygun & Adadan, 2019; Wen et al., 2018). Therefore, instructional designs that scaffold students in using the virtual labs for scientific inquiry become critical. Among diverse formats of instructional designs, teacher demonstration or modeling (Collins, 2006) and student critique (Chang & Linn, 2013) represent integral cognitive apprenticeship approaches to skill development in the contextualized settings. This study thus investigated the impact of two instructional designs, teacher demonstration and student critique, as the scaffolding for the development of scientific literacy to augment the effect of using virtual labs.

The affordance of virtual labs

Virtual labs provide an interactive space for students to explore the scientific concepts embedded in science phenomena (Heradio et al., 2016; van Joolingen et al., 2005), and thus allow learners to conduct virtual experiments and enable them to engage in core scientific practices, especially on phenomena that cannot be easily observed or investigated in real-life situations. Recently, many high-quality virtual labs have been made available for teachers and students on online platforms such as PhET, Molecular Workbench, Go-Lab, and CoSci, which offer the opportunity for students to manipulate a specific science context. Extensive studies have shown the affordances of the virtual labs in supporting science

learning (Chang et al., 2020; Vergara et al., 2017). The study by Martinez et al. (2011) found that students using virtual labs and physical laboratories achieved similar levels of conceptual understanding. In the study by Nolen and Koretsky (2018), they found that students demonstrated a higher level of motivation when they used virtual labs. Participating in an inquiry activity with virtual labs helped students to improve their experiment design skills (Lefkos et al., 2011).

Physical labs, on the contrary, have unique features that cannot be afforded by virtual labs. For instance, physical labs provide a feeling of realism (Abdulwahed & Nagy, 2011) and have the advantage of supporting experiments that require kinesthetic manipulation and tactual sensation (Zacharias et al., 2008), and thus provide a bridge from concrete to abstract conceptualization. Both virtual and physical labs have their own unique affordances for supporting science learning. However, virtual labs visualize abstract science concepts with multiple representations, and thus help students link new knowledge and previous knowledge (Taramopoulos & Psillos, 2017). Despite the positive affordances of the virtual labs, it should be noted that the inquiry based on these virtual labs is still challenging for students, as students tend to interact with virtual labs simply at a superficial and playful level (Swaak & de Jong, 2001). Therefore, scaffolding is needed to help students achieve in-depth understanding of science concepts through the use of virtual labs.

An integral affordance of virtual labs is their openness. This openness allows teachers to adopt an open-ended inquiry approach that enables students to take control of their own experiments to conduct science inquiry (Wen et al., 2020). Efstathiou et al. (2018) and Chang et al. (2020), for instance, engaged their students in using virtual labs to conduct virtual experiments to learn science concepts related to sinking, floating, and relative density. These studies suggested that with the guidance provided by the virtual labs, students achieved a higher level of general inquiry skills, including the ability to identify variables, state hypotheses, and operationally define and design investigations across various contexts, than their counterparts. However, another study (Wen et al., 2018) compared the strategies adopted by students who successfully used a virtual lab to construct scientific models and those who were not successful. The authors found that half of the students could not build a scientific model with the virtual lab as they had difficulties linking the scientific concepts with the virtual lab context.

The study above suggested that simply using virtual labs does not guarantee the development of scientific literacy. Virtual labs often help students visualize the processes, but students may not necessarily change their mental models after viewing them (Akaygun & Adadan, 2019). It is stressed that students have difficulties conducting mindful and purposeful virtual experiments (McElhaney et al., 2015). Teaching guidance is therefore needed to support learners' inquiry with virtual labs to conduct scientific virtual experiments (Efstathiou et al., 2018; Thoms & Girwidz, 2017).

Scaffolding embedded in virtual labs

One of the challenges that students face in the development of scientific literacy is that they need to think and act like scientists who must be sensitively alert to methodological flaws and various sources of error (Charney et al., 2007). The canonical science involves subtle associations among background information, data, claims, counterclaims, and rebuttals of authentic science. Therefore, the development of scientific literacy requires students to deliberately practice and reflect on the complex associations so that they can think and act like scientists (Yore & Treagust, 2006). However, recent science teaching practice in

classrooms places more emphasis on students' ability to recognize the standard scientific explanation, but less on helping them think why an answer might be flawed, what is wrong with an experimental design, why an interpretation of a dataset might be incorrect, or how to improve a weak explanation (Henderson et al., 2015).

The 'Learning Science by Doing Science' (LSDS) framework (Labouta et al., 2018) has been frequently applied to cultivate students' scientific literacy. The framework includes integral principles including student-centered, self-directed learning, and skills-based learning outcomes. In this line of research, diverse types of scaffolding are combined with virtual labs to provide different levels of guidance during different stages of inquiry with the virtual labs. Science inquiry involves complex processes that need students to apply scientific strategies to control the investigation. Scaffolding may be provided in diverse forms that guide students to productively use the virtual lab before, during, or after the inquiry with the virtual lab. Previous studies have coined several theory-informed frameworks that operationalize scaffoldings in virtual labs (see Quintana et al., 2004 for a summary). On one hand, virtual labs can be designed with proper representations to help students make sense of the target science phenomena (e.g., Chao et al., 2016; Wu et al., 2002). Furthermore, the virtual lab can explicitly organize the inquiry structure and make it clear to students to help them model the science process. For instance, the virtual lab in the study by Wen et al. (2020), allowed students to create their own inquiry map that guides the inquiry process. Their findings suggested that the inquiry with guidance embedded in the virtual lab had a long-term effect on the students' scientific literacy.

On the other hand, virtual labs can explicitly promote scientific reasoning strategies through prompting questions. This format of scaffolding involves proactive provision of contextualized prompts according to students' status in the virtual lab. With the contextualized prompts, students are more likely to focus on the important aspects of the science problem (Hmelo & Day, 1999). In particular, when they make problematic moves in their inquiry, a highly proactive prompt, that is, problematizing scaffolding, can be given to students to draw their attention to these problematic moves. A study by Efstathiou et al. (2018) indicated that such problematizing scaffolding can improve students' inquiry skills. Similarly, in a study by Li et al. (2019), adaptive scaffolds that asked students critical questions about the inquiry process were provided based on students' inquiry moves. Li et al.'s study confirmed that with the reflective questions provided by the virtual lab, students were better able to learn the science practices and transfer these competencies to new topics.

The above formats of guidance were system-dependent scaffolding, which means that they can be provided through the software system itself. Implicit scaffolding (Moore et al., 2016) is thus able to be provided by the use of constraints of the virtual lab systems to cue and guide students to engage in productive interactions. However, the system-dependent scaffolding is tightly integrated with the system while students are using the virtual lab. Teachers' ideas about using the virtual labs may not be compatible with the original design.

System-independent scaffolding

Another type of scaffolding is system-independent scaffolding that teachers may apply while considering the virtual labs' affordance and limitations. Such scaffolding incorporates teachers' idea about how to use the virtual lab before, during, or after students' engagement in the learning with the virtual labs. System-independent scaffolding is crucial in promoting scientific literacy as teachers may have different ideas about using existing virtual labs which cannot be customized to implement these ideas. In this track of

research, instructional approaches may be implemented to augment the effect of the virtual labs. For instance, the light-weight approach that guides students to use virtual labs with paper worksheets may feasibly promote students' scientific literacy (Chang et al., 2020). The study by Authors (Chang, 2017) indicated that driving questions given by teachers may guide students to conduct meaningful experiments and thus result in better learning efficiency compared with students who received the teacher's structured process prompts, suggesting that different formats of instructional approaches to using virtual labs may lead to different levels of learning effects.

The cognitive apprenticeship perspective may shed light on a holistic picture of such instructional approaches to developing scientific literacy with virtual labs. Cognitive apprenticeship asserts that skills are instrumental to the accomplishment of meaningful real tasks in practice, and emphasizes generalizing knowledge so that it can be used in many different settings (Collins, 2006). Diverse cognitive apprenticeship principles can be applied to develop scientific literacy in order to carry out real science inquiry practices. Scaffolding should be provided to guide students to carry out the essential tasks in a domain. Furthermore, cognitive apprenticeship particularly addresses the need to help students build strategic knowledge when carrying out these tasks. Modeling, for instance, is an integral pedagogical approach of cognitive apprenticeship that involves an expert demonstrating a task so that students can observe and build a model of the task (Larkins et al., 2013; Liu, 2005; Oriol et al., 2010). Another approach to cognitive apprenticeship is articulation that asks students to explicitly state their knowledge and reasoning when carrying out a task (Bouta & Paraskeva, 2013; Larkins et al., 2013).

Teacher demonstration and student critique were frequently applied as instructional approaches to cognitive apprenticeship in science classrooms since scientific literacy involves both explicit knowledge and implicit strategies in doing science. Teacher demonstration involves an expert demonstrating a task so that students can observe and build a model of the task (Larkins et al., 2013; Liu, 2005; Oriol et al., 2010). A study by Sever et al. (2010) indicated that teacher demonstration of scientific experiments can enhance students' learning performance in the dimensions of analysis, evaluation, and synthesis. The inclusion of teacher demonstration before student engagement in inquiry with virtual labs may help students model the inquiry process with virtual labs.

Critiquing is another of the applicable activities that involve students in playing a critical role of expressing their understanding of a task. Henderson et al. (2015) indicated that overlooking the role of critiquing in science education may lead to a failure to develop students' analytical reasoning ability. Research has found that engaging students in the practice of critiquing can help them develop integrated understanding of science concepts (Chang & Chang, 2013; Chang & Linn, 2013; Akaygun & Adadan, 2019), essay writing (Mørch et al., 2017), mathematical formulation (Wilkie & Ayalon, 2019), and scientific explanation skills (Matuk et al., 2019). It has been suggested that it is critical for students to learn to critique inquiry processes, as scientists must be alert to methodological flaws and various sources of error (Charney et al., 2007). Therefore, the critique approach may be helpful in building students' strategic knowledge of executing a task.

Both the teacher demonstration and student critiquing approaches have the potential to enhance the effect of virtual labs due to their shared features including connections to concrete examples and strategies to accomplish a cognitive task. However, they are different in their features of supporting learning including the concreteness of the examples and cognitive difficulties, and thus the two instructional approaches have certain strengths and weaknesses in supporting the learning of scientific inquiry. It has been indicated that even with teachers' demonstration in mind, students produced

a significant number of flawed inquiry cases (Wen et al., 2020), and students' inquiry plans were still imprecise and ill-defined after five guided-inquiry laboratory tasks (Crujeiras-Pérez & Jiménez-Aleixandre, 2017). On the other hand, the literature also indicates that students have difficulty critiquing scientific artefacts (Lai et al., 2016). Although the critique strategy has been applied to support science learning, the critique activity mainly targets the quality of learning outcomes such as the models built by students rather than the process of inquiry (Chang & Chang, 2013). Whether students can productively detect the flaws associated with the inquiry process is still not clear in the literature.

The present study

This study is part of a 3-year project aiming to enhance students' scientific literacy through virtual labs that are available on the Internet. The project is based on the rationale that virtual labs as a stand-alone application cannot guarantee effective learning (Renken & Nunez, 2013) and little is yet known about effective instructional approaches which could be combined with virtual labs to augment their effectiveness. This study thus worked with two science teachers at the participating school based on the LSDS framework in which the teachers engaged students in the inquiry with virtual labs. However, the target participants of our project were middle school students who needed extra scaffolding before engaging in science inquiry with the virtual lab. This study thus designed two instructional approaches, that is, teacher demonstration and student critique, based on cognitive apprenticeship, as the main schemes to prepare students with skills to productively use the virtual lab.

The two approaches aimed to strengthen students' understanding of the scientific process and judgement rather than the science knowledge before they took part in the inquiry. Previous studies have indicated the close relation between students' performance in inquiry and their understanding of the science process (Karamustafaoğlu, 2011; Wen et al., 2020). In other words, if students develop higher order process skills through authentic inquiry experiences, they are more likely to engage in science inquiry in an effective and scientific manner (Roth & Roychoudhury, 1993). However, how students' performance in the instructional activity before the inquiry relates to their inquiry performance and the scientific literacy after the whole process is not clear in the literature.

This study thus investigated the effects of the teacher demonstration and student critique and the relations among students' inquiry performance with the virtual lab, critique performance, and their scientific literacy. Multiple data sources were analyzed to answer the following research questions:

RQ1 What is the effect of the teacher demonstration and student critique approaches on students' scientific literacy?

RQ2 What is the effect of the teacher demonstration and student critique approaches on students' inquiry activities with the virtual lab?

RQ3 What is the relation, if any, between students' inquiry performance with the virtual lab and their scientific literacy?

RQ4 What is the relation, if any, among the students' critique performance, inquiry performance with the virtual lab, and their scientific literacy?

Method

Participants

This study aimed to understand how students learn science and develop scientific literacy with virtual labs under different instructional designs. The participants were 50 eighth-grade students, aged 14 to 15 years, from two intact classes at a middle school in northern Taiwan. The participants at this age were science inquiry novices. The two classes were randomly assigned to one of the treatments: teacher demonstration followed by student inquiry with a virtual lab (the teacher demonstration group, $n=24$) or student critique followed by student inquiry with a virtual lab (the student critique group, $n=26$). The two classes involved a virtual lab that enabled students to take part in the science inquiry practice. The target topic of learning was buoyancy. The two groups of students did not differ in terms of their prior scientific literacy of the topic of buoyancy as measured by the pre-tests ($t(47)=0.33, p=0.75$), indicating that they had similar prior knowledge.

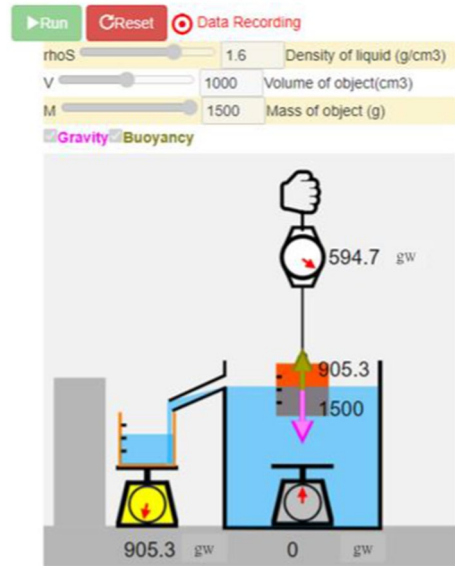
Procedure

This study compared students' scientific literacy and inquiry performance after they participated in guided inquiry, that is, teacher demonstration and student critiquing, which represents different instructional approaches to fostering scientific literacy. The students of the two groups went through three main phases: a pre-test, guided inquiry activities with a virtual lab, and a post-test. Each individual student of the two groups took pre-tests before and post-tests after the treatments to identify any possible changes in their scientific literacy, and students' inquiry activities with the virtual lab were analyzed to understand their inquiry performance. Both the pre-test and post-test lasted about one class period (45 min). After the pre-test, the two groups participated in the guided inquiry activity to learn the concept of buoyancy also for three class periods.

The students of the teacher demonstration and student critique groups were guided to conduct science inquiry in a guided inquiry map system (described later) that enabled them to participate in meaningful science activities with a virtual lab. However, students as novices of science inquiry needed guidance to understand the essential tasks and critical strategies to make scientific decisions. Therefore, in the teacher demonstration treatment, the teacher spent one class period (45 min) demonstrating how to conduct experiments with a buoyancy virtual lab (Fig. 1) provided on the CoSci platform (<https://cosci.tw/>). The buoyancy virtual lab visualizes the buoyancy of an object hanging on a hand-held scale. Students could manipulate the density of the liquid, the volume of the object, and the mass of the object to observe the relationship between these variables. The readings of the three scales can also help students understand the relationship among the buoyancy, the weight of the object, and the weight of the liquid replaced by the object.

The teacher of the teacher demonstration group selected one inquiry question provided on the CoSci platform and showed her students how to formulate hypotheses, design and conduct virtual experiments, collect and analyze data, and make

Fig. 1 The buoyancy virtual lab



conclusions to address the inquiry question. All materials needed for the modeling were prepared in advance. During the demonstration activity, the teacher showed how inquiry is conducted with the virtual lab, and occasionally paused to ask students questions to engage them in mindful observation. The students were then allowed to conduct their own inquiry with the guided inquiry map system for two class periods (90 min).

The student critique treatment asked students who were novices of science inquiry to critique flawed inquiry instances before they conducted science inquiry. The design of the student critique activities was based on Chang and Linn (2013) who indicated that critiques involving reflection facilitate knowledge integration. Instead of watching the teacher's demonstration, the students of the critique group worked on worksheets prepared by the teacher that asked them to critique fictitious experiment instances with the buoyancy virtual lab. For example, the students were given an inquiry question and a series of experiment designs and were asked to critique "Whether these designs can answer the inquiry question or not. If not, how can you improve the designs?" The students were also asked to critique whether a given set of data could be used to support a given conclusion, and how to improve the conclusion. The students spent one class period (45 min) completing the critique worksheets. The teacher also led whole class discussions to engage the students in discussing their critiques. Then the students conducted their own inquiry with the guided inquiry map for two class periods. It should be noted that the virtual labs used by the student critique and teacher demonstration groups were identical, and they manipulated the same set of variables.

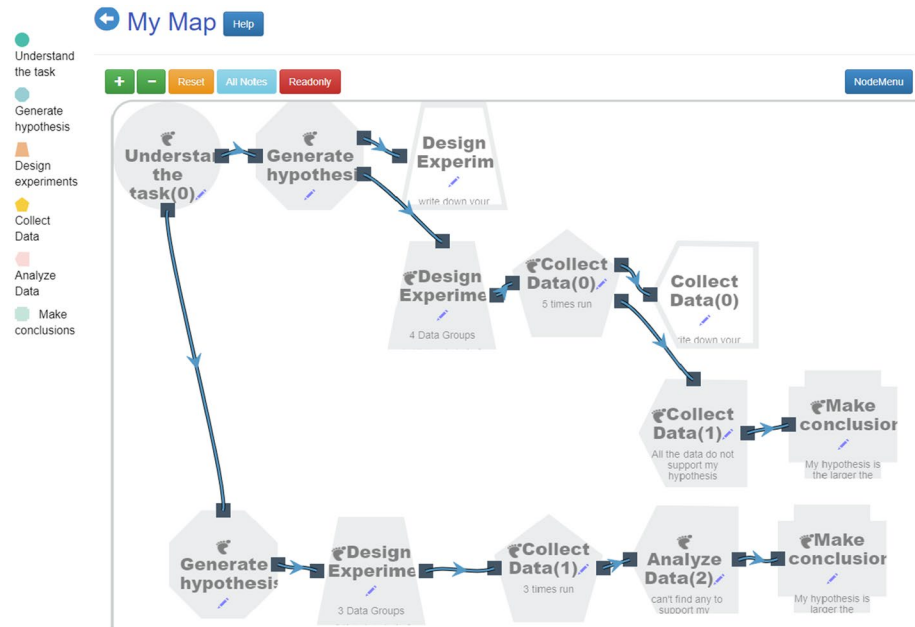


Fig. 2 The guided inquiry map system

Guided inquiry map system

The teacher demonstration and student critique groups participated in a science inquiry activity to showcase their processes of building the target concepts of buoyancy. To support students in building the target concepts, this study developed a guided inquiry map system (Fig. 2). The system is an inquiry component of the CoSci platform which provides over 100 free science virtual labs, and allows students to conduct inquiries using the virtual labs on the platform. The guided inquiry map system serves as an activity design system for teachers and an inquiry system for students. With the support of the CoSci platform, the teachers can design an inquiry activity by selecting an available virtual lab and specifying the goal of the inquiry with the virtual lab. It also allows teachers to provide candidate hypotheses for the virtual lab if the concepts associated with the hypotheses are critical, and if students have problems with the generation of hypotheses related to the virtual lab. Since the time for the inquiry activity is limited, the teacher provided six predefined hypotheses that can be verified by the students with the virtual lab in Fig. 1 including:

H1 The heavier an object is, the more likely it is that this object will sink.

H2 The bigger an object is, the more likely it is that this object will sink.

H3 The buoyant force is equal to the weight of the fluid displaced by the object.

H4 As an object sinks, the buoyant force keeps increasing after it is completely immersed in the fluid.

H5 The buoyant force is equal to the object’s volume under the fluid times the density of the fluid.

H6 Whether an object can float or sink depends on the relative density of the objects to the fluid.

These hypotheses may involve different levels of complexity of scientific inquiry. The students of the two groups could freely select the hypothesis that they wanted to investigate. Therefore, the hypotheses that the two groups selected may also be an indicator of the impact of the treatments.

Design Experiment(0)🔍

In design experiment phase, scientists design appropriate parameter sets to verify their hypothesis. Scientists conduct the experiment based on the design, observe the process and collect data.

Please design the experiment according to your hypothesis. You can click the 'Add' button to create new datasets, and fill valid number into it. Click the 'send' button for recording the datasets.

rhoS - density of the liquid (g/cm3) range : 0.6 ~ 2.0
 M - mass of object(g) range: 500 ~ 1500
 V - volume of object(cm3) range: 500 ~ 1500

Design1 rhoS: 1 M: 1000 V: 800	Design 2 rhoS: 1 M: 600 V: 800	Design 3 rhoS: 1 M: 800 V: 800	Design 4 rhoS: 1 M: 1200 V: 800
---	---	---	--

Collect Data(0)🔍

Using the simulation to conduct experiment based on your experiment design, observe and collect the data. (You can click the box below to set the simulation with the number in the box.)

Design1 rhoS: 1 M: 1000 V: 800	Design2 rhoS: 1 M: 600 V: 800	Design3 rhoS: 1 M: 800 V: 800	Design4 rhoS: 1 M: 1200 V: 800
---	--	--	---

The screenshot shows a simulation interface with control sliders on the left and a variables table on the right. The sliders are set to rhoS: 1, M: 1200, and V: 800. The variables table lists various parameters and their values.

Variable	Description	Value	Unit
t	time	0.00	sec
rhoS	density of liquid	1.00	g/cm3
M	mass of object	1200.00	g
V	volume of object	800.00	cm3
rhoA	density of object	1.50	g/cm3
Vu	volume of object under liquid	0.00	cm3
Vs	Volume of liquid leaking out	0.00	cm3
scale1	Grey scale	0.00	gw
scale2	Yellow scale	0.00	gw
scale3	Hand held scale	1200.00	gw

Fig. 3 The “design experiment” and “collect data” nodes

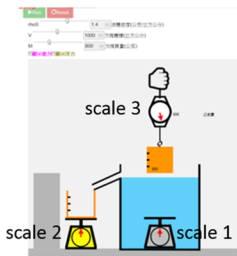
Students of this study conducted their inquiry associated with the buoyancy with the guided inquiry map system. The system provided the students with structural guidance for the inquiry process to enhance their process awareness of the science inquiry practice. To begin an inquiry thread, students could select any of the predefined hypotheses or generate their own hypotheses in the "generate hypothesis" node. After specifying their hypothesis, a "design experiment" node would be created which guided the students to specify a set of variable settings to verify their hypothesis (the left in Fig. 3). Then they could proceed to collect the data in the "collect data" node which was created after the students completed the experiment design (Fig. 3). Upon clicking the "collect data" node, the buoyancy virtual lab together with the variable settings predefined by the students was provided for the students to conduct the experiment. The data of the virtual lab results were automatically recorded and shown in the "analyze data" node through which students could compare the results of the virtual lab under different variable settings. At the end of the inquiry thread, a "make conclusion" node was displayed for the students to conclude whether their hypothesis was supported or rejected by the experiment results. The system allowed students to conduct multiple threads of inquiry and displayed all the inquiry threads in a node and link structure. Therefore, the system functioned as structural scaffolding for students to understand the status of their inquiry activity as it provided structural supports for students to plan, execute, and reflect on their inquiry tasks. All the activities during the inquiry process were logged for the investigation of the quality of the students' inquiry activities with the virtual lab.

The teacher of the teacher demonstration group demonstrated the inquiry process for the scientific hypothesis H2 with the inquiry map system. She prepared presentation slides to outline the screenshots about how she went through the "design experiment", "collect data", "analyze data", and "make conclusion" in the inquiry map system. In each screenshot for an inquiry phase, she presented the purpose of this phase and outlined her operations in the inquiry map system. After presenting the screenshots, the teacher operated the inquiry map system to showcase the real operations of the inquiry. Therefore, students understood the detail operations of the inquiry to address the scientific hypothesis H2.

Student critique worksheet

The students of the student critique group engaged in a student critiquing activity before taking part in the inquiry activity. A student critique worksheet was developed to help the students engage in the critiquing activity. The worksheet was designed based on the problematic inquiry instances that participants of a prior study frequently demonstrated (Wen et al., 2020). These problematic inquiry instances were frames to seven critique questions to examine whether the students could identify problems associated with the experiment designs (question-design) or the selection of experiment trials (question-trial) to verify a hypothesis and detect unreasonable conclusions from a given data collection (evidence-conclusion). The critique worksheet is shown in Appendix 2. The critique questions were designed based on the scenario of the inquiry using the virtual lab that the teacher used for demonstration for the teacher demonstration group. In other words, the critique and the teacher demonstration activities envisioned similar pictures of the inquiry with the virtual lab. Therefore, they can understand the basic process of the inquiry with the virtual lab. Figure 4 displays an example question from the worksheet asking whether a conclusion is reasonable (evidence-conclusion) with the data provided by the virtual lab.

	ρ_{obj}	M_{obj}	V_{obj}	ρ_{fluid}	V_{fluid}	V_{sub}	scale1	scale2	scale3	Buoyant F_{b}
Exp1	1.00	600.00	1000.00	0.60	600.00	600.00	0.00	600.00	0.00	600
Exp2	1.00	800.00	1000.00	0.80	800.00	800.00	0.00	800.00	0.00	800
Exp3	1.00	1000.00	1000.00	1.00	1000.00	1000.00	0.00	1000.00	0.00	1000
Exp4	1.00	1200.00	1000.00	1.20	1000.00	1000.00	0.00	1000.00	200.00	1000
Exp5	1.40	1400.00	1400.00	1.00	1000.00	1000.00	0.00	1400.00	0.00	1400



According to the results above, Hsian-Hsian claims that “I hypothesized that the buoyant force equal to the weight of the fluid displaced by the object. The experiment results support my hypothesis. So, I will not change my hypothesis. This is because the weight of the fluid displaced by the object is equal to the volume of the object under the fluid. The volume of the object under the fluid is equal to the buoyant force received by the object. So the weight of the fluid displaced by the object is equal to the volume of the object under the fluid.” Is this claim reasonable, if not, how do you revise the claim?

Fig. 4 An example of a critique question on the critique worksheet

Data collection and analysis

Pre-test and post-test

This study applied the test developed by Authors (Wen et al., 2020) to understand students’ scientific literacy in the context of buoyancy before and after the treatment. The questions are shown in Appendix 1. The test involves eight questions related to buoyancy. The eight items measure four dimensions of scientific literacy based on the OECD scientific literacy framework (OECD, 2016) since these dimensions are directly related to the inquiry activity facilitated by the use of virtual labs and the guided inquiry map. The four dimensions include important skills that are needed in inquiry covering both procedural knowledge and scientific reasoning skills including offering explanatory hypotheses (SC-H), identifying the question explored in a given scientific study (SC-Q), interpreting data and drawing appropriate conclusions (SC-C), and evaluating ways of exploring a given question scientifically (SC-E). The items went through several rounds of revision by two science educators, two science education researchers and one assessment expert to ensure content and construct validity.

Students’ responses to the pre-test and post-test questions were evaluated by two independent coders. Since the questions include a claim part and a reason part, detailed scoring rubrics for the claim and reason parts were developed and used by the two coders. A student’s claim part was given 1 point if it was consistent with the scientific knowledge, otherwise it would be given 0 points. An answer to the reason part would be given 2 points if it completely and scientifically explained the reason of the claim part, 1 point for a partial but scientific reason, and 0 for an irrelevant explanation. The two coders independently scored the tests, and inconsistent codes were discussed and resolved. The two independent coders coded 20% of the tests following the rubrics, and the inter-coder agreement reached 0.94 (Cohen’s Kappa) showing high reliability of the coding (Lombard et al., 2002). A pre- and post-test comparison was made to identify whether there was any enhancement of literacy

within each group. Furthermore, ANCOVA was employed to test the differences among the two treatments, using the post-test score as the dependent variable, the pre-test score as the covariate, and the treatment as the independent variable.

Students' inquiry performance in the virtual lab

All the activities students performed in the guided inquiry map system were logged and analyzed to understand their engagement in the inquiry activity. The analysis involved a sequence of analyses from quantity to quality analyses to obtain students' inquiry performance. First, the number of nodes created in the guided inquiry map and the number of executions of different types of inquiry tasks including the hypothesis generation, experiment design, data collection, data analysis, and conclusion were analyzed. Second, the number of distinct inquiry moves were counted because some of the inquiry nodes represent similar inquiry moves to those of other nodes or did not involve meaningful inquiry moves. Third, the final inquiry performances were obtained from students' responses to the prompting questions in each of the distinct inquiry nodes. The responses were analyzed based on the scoring rubric developed in a study by Authors (Wen et al., 2020) (as shown in Table 1). Two coders scored all the prompts generated by the students in the teacher demonstration and critique groups. Students' scores in each inquiry phase were summarized to obtain a score reflecting their inquiry performance in each phase. The reliability of coding the students' quality of inquiry process was 0.91 (Cohen's Kappa) indicating that the coding is highly reliable. The qualitative and quantitative data of the teacher demonstration and critique groups were compared with the *t* test to understand the differences in their behavioral engagement and inquiry performances.

Student critique performance

Two coders evaluated the critique group's answers on their critique worksheet. The evaluation considered whether they correctly identified the problematic methodological designs, analyses, and inferences, and their explanation of why they were problematic. Similar to the evaluation of the scientific literacy test, 2 points were given for a completely adequate response, 1 point for a partially adequate response, and 0 for an incorrect one. The two independent coders coded 20% of the critique worksheets following the rubrics, and the inter-coder agreement reached 0.91 (Cohen's Kappa) showing the high reliability of the coding.

Results

The scientific literacy test (RQ1)

The students' pre-test mean scores and standard deviations are summarized in Table 2. The independent *t*-test analysis indicates that there was no significant difference between the two groups for overall scientific literacy before the treatments ($t(47)=0.75$, $p=0.33$). The two groups of students also did not show significant differences in any of the detailed scientific literacy dimensions including offering explanatory hypotheses (SC-H) ($t(47)=1.40$, $p=0.17$), identifying the question explored in a given scientific study (SC-Q) ($t(47)=-0.33$, $p=0.74$), interpreting data and drawing appropriate conclusions (SC-C)

Table 1 The scoring rubric for the quality of the students' inquiry process (Wen et al., 2020)

Inquiry phase	Scoring Rubric			
	Score 0	Score 1	Score 2	
Experiment design	<p>Experiment design A – Student designed experiments aligned with the selected inquiry question</p> <p>Experiment design B – Student designed experiments based on the experiment principle</p>	<p>Irrelevant description, blank, or not aligned with the inquiry question</p> <p>Design not based on the experiment principle, namely, varying one variable at a time</p>	<p>Description of the designs align with the inquiry question</p> <p>Designs based on the experiment principle</p>	<p>N/A</p> <p>N/A</p> <p>N/A</p>
Data collection and analysis	<p>Conducting trials – Student conducted trials of experiments aligned with the selected inquiry question</p> <p>Data analysis – Student identified variables needed to be analyzed in order to address the inquiry question</p>	<p>Irrelevant description, blank, or trials not aligned with the inquiry question</p> <p>Irrelevant description or blank</p>	<p>Correct description of what data were selected from the trials that were used to address the inquiry question</p> <p>Identification of only part of the variables related to the inquiry question</p>	<p>N/A</p> <p>N/A</p> <p>Identification of all variables related to the inquiry question</p>
Conclusions	<p>Conclusions – Student made conclusions based on the evidence</p>	<p>Irrelevant description or blank conclusions based on the evidence</p>	<p>Conclusion made consistent with the scientific principle of buoyancy</p>	<p>Conclusion made consistent with the scientific principle of buoyancy and data used as evidence</p>

Table 2 The independent *t* test analysis of the pre-test of the science literacy score between the teacher demonstration and critique groups

Scientific literacy	Group	<i>N</i>	<i>M</i>	<i>S.D</i>	<i>t</i>	<i>P</i>
Pre Overall	Critique	26	9.23	3.64	0.33	0.75
	Demonstration	23	8.87	4.13		
Pre SC-H	Critique	26	2.38	1.33	1.40	0.17
	Demonstration	23	1.74	1.89		
Pre SC-Q	Critique	26	0.50	0.71	- 0.33	0.74
	Demonstration	23	0.57	0.66		
Pre SC-C	Critique	26	6.08	2.35	- 0.22	0.83
	Demonstration	23	6.22	2.09		
Pre SC-E	Critique	26	0.27	0.53	- 0.54	0.59
	Demonstration	23	0.35	0.49		

($t(47) = -0.22, p = 0.83$), and evaluating ways of exploring a given question scientifically (SC-E) ($t(47) = -0.54, p = 0.59$). Therefore, the students of the two groups did not show significant differences in their scientific literacy before the treatments.

The results of the dependent *t*-test of the pre- and post-test for the teacher demonstration and student critique group are shown in Table 3. Overall, the two groups demonstrated significant improvement in their scientific literacy after participating in the learning activity. The two groups both improved their scientific literacy in the dimension of SC-C (interpreting data and drawing appropriate conclusions). Significant improvement in the SC-Q dimension (identifying the question explored in a given scientific study) and SC-E (evaluating ways of exploring a given question scientifically) were found in the student critique group.

The ANCOVA analysis of the students' post-test scores using the pre-test scores as the covariate are summarized in Table 4. The ANCOVA results indicate that there was a significant treatment effect on the overall scientific literacy score ($F(1,46) = 4.20, p < 0.05$). Furthermore, the analysis found a significant difference between the student critique and teacher demonstration groups. The students in the critique group demonstrated a higher level of improvement in evaluating ways of exploring a given question scientifically (SC-E)

Table 3 The learning gains of science literacy for the teacher demonstration and student critique groups

		Pre-test		Post-test		<i>t</i>	<i>p</i>
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Overall	Critique	9.23	3.64	14.15	6.03	- 7.32**	<0.01
	Demo	8.87	4.13	11.65	3.88	- 3.43**	<0.01
SC-H	Critique	2.38	1.33	2.58	1.70	- 0.68	0.50
	Demo	1.74	1.89	1.65	1.27	0.27	0.79
SC-Q	Critique	0.50	0.71	1.08	1.23	- 0.44*	0.02
	Demo	0.57	0.66	0.61	0.89	- 0.19	0.85
SC-C	Critique	6.08	2.35	8.04	3.00	- 3.64**	<0.01
	Demo	6.22	2.09	7.17	1.88	- 2.21*	0.04
SC-E	Critique	0.27	0.53	0.69	0.74	- 2.85**	<0.01
	Demo	0.35	0.49	0.26	0.45	0.81	0.43

* < 0.05, ** < 0.01

Table 4 The ANCOVA analysis of the science literacy between the teacher demonstration and student critique groups

Dimensions	Sources	Type III sum of squares	<i>df</i>	Mean square	<i>F</i>	<i>p</i>	Partial η^2
SC Total	Pre science literacy total (Covariate)	615.87	1	615.87	45.35	0.00	0.50
	Group	57.03	1	57.03	4.20	0.05*	0.08
	Error	624.73	46	13.58			
SC-H	Pre SC-H (Covariate)	32.53	1	32.53	19.94	0.00	0.30
	Group	4.11	1	4.11	2.52	0.12	0.05
	Error	75.04	46	1.63			
SC-Q	Pre SC-Q (Covariate)	2.27	1	2.27	1.97	0.17	0.04
	Group	2.91	1	2.91	2.53	0.12	0.05
	Error	53.06	46	1.15			
SC-C	Pre SC-C (Covariate)	67.97	1	67.97	13.35	0.00	0.23
	Group	10.78	1	10.78	2.12	0.15	0.04
	Error	234.29	46	5.09			
SC-E	Pre SC-E (Covariate)	2.08	1	2.08	6.03	0.02	0.12
	Group	2.61	1	2.61	7.55	0.01**	0.14
	Error	15.89	46	0.35			

* <0.05 , ** <0.01

($F(1,46)=7.55, p<0.01$). Overall, the results indicate that the student critique group demonstrated a higher level of improvement in their scientific literacy than did the teacher demonstration group.

Students' inquiry performance in the virtual lab (RQ2)

To understand whether the critique and teacher demonstration group treatments influenced the student inquiry process, this study analyzed their inquiry process on the guided inquiry map system. As shown in Table 5, the critique group created an average of 12.15 nodes, while the teacher demonstration group created 12.83 nodes. While the teacher

Table 5 Comparison of numbers of nodes generated in the virtual lab

	Critique ($n=26$)		Demonstration ($n=24$)		<i>t</i>	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Hypothesis node	2.27	0.67	2.50	0.72	-1.18	0.25
Experiment design node	2.77	1.53	2.71	1.46	0.14	0.89
Data collection node	2.69	1.29	2.92	1.35	-0.60	0.55
Data analysis node	2.31	1.01	2.58	0.88	-1.02	0.31
Conclusion node	2.12	0.71	2.13	0.85	-0.04	0.97
Total node	12.15	3.82	12.83	4.34	-0.59	0.56

Table 6 Comparison of distinct inquiry moves in the guided inquiry map

	Critique (<i>n</i> = 26)		Demonstration (<i>n</i> = 24)		<i>t</i>	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Number of hypotheses generated	2.27	1.12	2.29	0.69	- 0.09	0.93
Number of experiments planned	8.23	3.80	11.17	4.14	- 2.62*	0.01
Number of trials of virtual labs	10.96	4.21	13.83	5.07	- 2.19*	0.03
Number of inquiry paths completed	1.81	0.80	2.08	0.65	- 1.33	0.19

**p* < 0.05

Table 7 Comparison of inquiry performance in the guided inquiry map

	Critique (<i>n</i> = 26)		Demonstration (<i>n</i> = 24)		<i>t</i>	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Experiment design A	1.19	0.69	1.58	0.97	- 1.64	0.11
Experiment design B	1.42	0.76	1.63	0.77	- 0.93	0.36
Trial conducting	0.58	0.81	0.50	1.02	0.30	0.77
Data analysis	1.31	0.97	2.58	1.28	- 4.00***	0.00
Conclusion	0.62	0.80	1.13	0.85	- 2.18*	0.03
Total	5.12	3.05	7.42	3.36	- 2.54*	0.01

p* < 0.05; **p* < 0.001

demonstration group created more nodes, there was no significant difference in the number of nodes created in the guided inquiry map by the two groups (*t*(48) = -0.59, *p* = 0.56).

Detailed analysis of the content of the nodes found that the teacher demonstration group seems to have demonstrated more distinct inquiry moves than the critique group did (Table 6). The teacher demonstration group completed more inquiry paths from the hypothesis, experiment design, data collection, data analysis to the conclusion than the critique group did. The teacher demonstration group completed an average of 2.08 paths, while the critique group planned only 1.81 paths. The difference between the two group was not significant (*t*(48) = - 1.33, *p* = 0.19). However, the teacher demonstration group completed an average of 13.83 trials of virtual labs which is significantly higher than the 10.96 that the critique group completed (*t*(48) = - 2.19, *p* = 0.03). A significant difference can also be found in the number of experiments planned. The teacher demonstration group planned 11.17 experiments, whereas the critique group planned only 8.23 experiments (*t*(48) = - 2.62, *p* = 0.01).

Students' responses in each of the distinct inquiry moves were evaluated and summarized to reflect their inquiry performance. The scores shown in Table 7 display the inquiry performance of the two groups of students. Overall, the teacher demonstration group received a score of 7.42 which is significantly higher than the critique group's score of 5.12 (*t*(48) = - 2.54, *p* = 0.01). In particular, the performances of the teacher demonstration group in the data analysis phase and conclusion phase were significantly higher than those of the critique group. This result is consistent with the results of the number of distinct

inquiry moves, indicating that the teacher demonstration group demonstrated more distinct inquiry moves than the critique group. As the inquiry performance is an accumulative score, the more trials of inquiry that the students conducted, the higher score they would receive in the inquiry performance.

This study further examined the hypotheses which students worked on in the guided inquiry map system. The examination of the hypothesis could help us understand what students intended to achieve when they were given different formats of instructions that can partially explain the cause of the differences in the inquiry performance between the teacher demonstration and critique groups. Both the demonstration group and the critique group could freely select target hypotheses from the candidate hypotheses (H1–H6) or specify their own hypotheses. It should be noted that the teacher demonstrated an example of inquiry using hypothesis H2 for the teacher demonstration group. Table 8 displays the number of students who worked on each hypothesis. The distribution shows that the teacher demonstration group demonstrated a higher level of engagement with hypothesis H1 which resembles H2 in their inquiry structure targeting whether a variable (volume and weight of an object) influences the state of the object. Twenty-three (46%) students worked on the two hypotheses. On the contrary, the critique group demonstrated a lower level of centrality on H1 and H2. Only 16 students (34%) worked on the two hypotheses. With the teacher's demonstration in mind, the teacher demonstration group tended to select and work on the hypotheses with similar structures, and thus were more proficient in conducting similar inquiry trials, accounting for the reason why they demonstrated more frequent behavioral engagement and higher scores in their inquiry performance.

The relation between scientific literacy and inquiry performance in the guided inquiry activities (RQ3)

Students' inquiry performance and inquiry moves in the guided inquiry map system were analyzed with their scientific literacy test scores to understand the roles of the guided inquiry activities in scientific literacy for the two different treatments. Table 9 shows the Pearson bivariate correlations among their scientific literacy scores and their inquiry performance. The results indicated that students' scientific literacy test scores were highly related to their number of distinct inquiry moves and inquiry performance in the guided inquiry map system under the critique treatment. More specifically, their pre-test scores of scientific literacy were significantly correlated to all the dimensions of inquiry (r ranges

Table 8 The number of students working on each hypothesis

	Critique	Demonstration
H1	1	22
H2*	15	1
H3	9	6
H4	9	7
H5	5	1
H6	6	8
Others	3	5
Total	48	50

*The teacher demonstrated an example inquiry process using this hypothesis in the teacher demonstration group

Table 9 Pearson bivariate correlation between the inquiry process and the scientific literacy test

		Critique (<i>n</i> = 26)		Demonstration (<i>n</i> = 23)	
		Pretest	Posttest	Pretest	Posttest
Number of nodes	Hypothesis	0.188	0.119	0.127	0.013
	Experiment design	0.383	0.329	0.140	0.221
	Data collection	0.349	0.423*	0.469*	0.178
	Data analysis	0.470*	0.281	0.344	0.243
	Conclusion	0.438*	0.462*	-0.059	0.258
Distinct inquiry moves	Number of hypotheses generated	0.369	0.214	0.120	0.032
	Number of experiments planned	0.341	0.393*	0.203	0.310
	Number of trials of virtual labs	0.432*	0.345	0.498*	0.359
Inquiry performance	Experiment design A	0.695**	0.709**	0.153	0.441*
	Experiment design B	0.573**	0.589**	0.048	0.359
	Trial conducting	0.443*	0.621**	0.080	0.149
	Data analysis	0.421*	0.395*	-0.083	0.035
	Conclusion	0.524**	0.574**	0.245	0.343
	Total	0.690**	0.749**	0.121	0.378

* < 0.05; ** < 0.01

from 0.421 to 0.695). Significant positive relations also exist between their post-test scores of scientific literacy and inquiry performance in the guided inquiry map system. Their post-test scores of scientific literacy were significantly correlated to the number of data collection nodes ($r=0.423$), the conclusion nodes ($r=0.462$), the number of experiments planned ($r=0.393$), and all the inquiry performance indicators (r ranges from 0.395 to 0.749). On the contrary, the correlations between the teacher demonstration group's scientific literacy test scores and their distinct inquiry moves and inquiry performance were not as obvious as those of the critique group. For example, the post-test scores of scientific literacy were not correlated with their inquiry performance except for experiment design A.

Such results indicate that students' inquiry activities with the virtual lab under the critique treatment align closely with the goal of scientific literacy as the critique activity informing the students of the general evaluation criteria of scientific inquiry. On the contrary, under the teacher demonstration treatment, the detailed demonstration example provided by the teacher may have interfered with the students' self-directed inquiry activities and attracted the students to work on similar inquiry hypotheses that implicitly restricted them from following the teacher's demonstration rather than exploring the target domain by themselves.

The relation between critiquing performance and students' inquiry performance/scientific literacy test (RQ4)

The critique groups' performance in the critiquing activity was analyzed together with their inquiry performance and their scientific literacy post-test score to display a clear picture of the role of the critiquing activity in the inquiry activity and their scientific literacy. Table 10 lists the correlations between the critique performance and these constructs. It

Table 10 Pearson bivariate correlation coefficients among critiquing performance, inquiry process, and the scientific literacy test ($n=26$)

	Question-design	Question-trial	Evidence-conclusion	Total
Number of nodes				
Hypothesis	- 0.045	0.174	0.028	- 0.032
Experiment design	0.184	0.030	- 0.147	0.032
Data collection	0.070	0.108	0.083	0.050
Data analysis	0.219	0.506**	0.137	0.352
Conclusion	0.185	0.294	0.235	0.015
Distinct inquiry moves				
Number of hypotheses generated	0.182	0.229	0.187	0.169
Number of experiments planned	0.151	0.147	- 0.012	- 0.087
Number of trials of virtual labs	0.217	0.326	0.159	0.177
Inquiry performance				
Experiment design A	0.567**	0.482*	0.304	0.497**
Experiment design B	0.652**	0.618**	0.392*	0.622**
Trial conducting	0.333	0.316	0.160	0.293
Data analysis	0.515**	0.448*	0.199	0.406*
Conclusions	0.357	0.405*	0.276	0.404*
Total	0.614**	0.574**	0.332	0.559**
Science Literacy Post Total	0.450*	0.486*	0.471*	0.436*

* < 0.05; ** < 0.01

was found that the critique performance is closely related to the scientific literacy post-test scores. The total critique performance is significantly correlated to the post-test score of scientific literacy ($r=0.436$). All sub-dimensions of the critique performance were also significantly correlated to the post-test score of scientific literacy. Furthermore, students' critique performances were also closely correlated to their inquiry performance in the virtual lab. More specifically, students' total critique performance is significantly correlated to their performance in experiment design A ($r=0.497$), experiment design B ($r=0.622$), data analysis ($r=0.406$), and conclusions ($r=0.404$). However, the critique performance is not closely related with the number of nodes created or with distinct inquiry moves. The results may be due to the fact that the critique activity involved students' qualitative evaluation of inquiry practice, and was thus closely linked to the quality of their inquiry process. The findings suggest that the students were able to transfer the implicit criteria they had learned in the critiquing activity to the self-directed inquiry activity.

Discussion and conclusions

The effect of the teacher demonstration and student critique on students' scientific literacy (RQ1)

An increasing number of studies have been conducted to understand the effect of virtual labs to enhance students' scientific literacy. In general, students can develop a higher level of scientific literacy when they learn with virtual labs than in traditional instruction environments (Ismail et al., 2016; Jannati et al., 2018; Quellmalz et al., 2020). It is also stressed that guided participation that helps students go through critical inquiry phases when they are using virtual labs is particularly important for improving scientific literacy and inquiry skills (Jin & Bierma, 2013). Ardianto and Rubini (2016) further asserted that the guided discovery pedagogical approach with virtual labs can achieve a similar level of improvement of scientific literacy to which students can achieve in the problem-based learning approach. The findings of this study echo this assertion, as the guided discovery based on the teacher demonstration and student critique pedagogical design was helpful for improving students' scientific literacy. As scientific literacy involves complex science processes and judgement in executing these process, the results of our study suggest that the critique activity used together with virtual labs is more influential than the teacher demonstration approach for enhancing student scientific literacy. Students who learned with the critique activity demonstrated significant improvement in identifying the question explored in a given scientific study and in evaluating ways of exploring a given question scientifically, while those students who learned with the teacher demonstration approach did not.

One possible explanation for the limitation of the teacher demonstration approach in supporting the development of scientific literacy in the classroom may be the discrepancy between the teacher's advanced mental model and the novice learners' model. As indicated by Järvelä (1995) "modeling, as performed by the teacher, hardly led to reciprocal understanding with the students. ...The cognitively more advanced mental model, which is different from and more complex than the students' own cannot be assimilated directly" (p. 256). This can partially explain the reason why the effect of the teacher demonstration was not as significant as that of the student critiquing approach. On the other hand, the critique activity prompted students to think deeply about the flaws involved in the inquiry process,

instead of passively observing the teacher's demonstration. Therefore, students could better link research questions and data and evaluate the ways to explore a scientific question.

The effect of the teacher demonstration and student critique pedagogies on students' inquiry activities (RQ2)

The teacher demonstration and student critique approaches demonstrated different impacts on inquiry activities with virtual labs. This study found that students in the teacher demonstration group demonstrated more distinct inquiry moves, and their total scores of inquiry performance were higher than those of the students in the critique group. As shown in the analysis of the hypotheses which the two groups worked on, the teacher demonstration group tended to select the hypotheses with a similar inquiry structure which the teacher demonstrated. Such a result reflects both the positive effect and the limitation of the demonstration approach. On one hand, the demonstrated inquiry instances acted as a worked example which the students could follow to regulate their inquiry actions. It in turn reduced the students' cognitive load in dealing with the complex inquiry process (Ginns et al., 2016; Sweller, 2005). This can explain the reason why the teacher demonstration group demonstrated more distinct inquiry moves and received higher scores in the inquiry performance.

However, the worked example also restricted the possibility for students to explore other scientific hypotheses and thus reduced the transferal effect by which students transferred what they had learned from the teacher's demonstration to new inquiry contexts. Such a phenomenon echoes the findings of previous studies by Dennen (2000) and Parker and Hess (2001) who indicated the limitation of teacher demonstration in the classroom. Demonstration can act as scaffolding in the form of chunking and sequencing the complex processes to guide students to appropriately perform the necessary tasks. However, their findings suggested that after the teacher demonstration, students tended to focus on the content being demonstrated, but with little emphasis on the method itself (Parker & Hess, 2001). In other words, the teacher demonstration enabled students to focus more on the content-based learning goals than on the project management of the learning tasks (Dennen, 2000). This limitation may result from the common constraint of the science classrooms, where one teacher has to face over 20 students, and thus demonstration may fail because the task and the situation are not reciprocally understood by the teacher and students (Järvelä, 1995).

The relationship between critique performance, inquiry activities, and scientific literacy (RQ3 and RQ4)

This study found that students' inquiry performance was significantly correlated with their scientific literacy scores in the student critique group. However, such close correlation was not found in the teacher demonstration group. In addition, the critique performance of the critique group also significantly correlated with their inquiry performance and their scientific literacy scores. Such results support that the critique activity is in alignment with the goal of developing students' inquiry skills and scientific literacy due to the fact that it involves students' qualitative evaluation of the inquiry practice, instead of focusing on the procedural knowledge which may be conveniently conveyed by teacher demonstration. It is such a counter-critique that requires students to identify flaws in a work and then show how the evidence in the counter-critique which demonstrated its weaknesses enhances students'

tendency to adopt skills (inquiry here in the present study) or perspectives that are considered important in the scientific community (Henderson et al., 2015).

The results of our study are consistent with those of the prior empirical studies. For instance, Namdar and Kucuk (2018) found that after participating in the critique activity, students were more likely to adopt an inquiry orientation perspective in their works. Wilkie and Ayalon's study also suggested that critiquing stimulated students' attention to the quality of science argumentation and helped them notice alternative pathways to view problems (Wilkie & Ayalon, 2019). Chang and Chang's (2013) study also confirmed the positive effect of critiquing on developing students' scientific modeling skills. However, critiquing is more difficult for students than construction of knowledge, and recent science teaching or learning practice has focused more on the ability to recall the standard scientific explanation, but rarely on the ability to discover what might be flawed (Henderson et al., 2015). It is thus necessary to understand how critiquing can be integrated into existing teaching and learning practices in current science education curricula to better develop such a critical competency.

Limitations and future research

The study demonstrated how the teacher demonstration and student critique approaches can be implemented with virtual labs in a regular science classroom. It demonstrated the positive impacts of critiquing activities on the learning of scientific literacy and how the critique approach can improve students' critical judgement of inquiry processes and their scientific literacy. This study also found the limitation of the teacher demonstration approach when it is combined with virtual labs. However, the findings were obtained based on the students' engagement in the learning with virtual labs. The findings may not be over-generalized to other settings such as the learning in a physical science lab as it involves more complicated inquiry settings than the virtual labs. Furthermore, the scaffolding involved in this study was provided before the students' engagement in the inquiry with the virtual lab. Further investigation is still needed to understand how other types of just-in-time scaffolding can be implemented in a more effective way to enhance the effectiveness of virtual labs. For instance, software agents that closely monitor students' progression in the guided inquiry system can be applied to provide just-in-time hints and questions to help students critically reflect on their inquiry process. Such an approach may be helpful to improve the effect of virtual labs. Furthermore, the participants of this study were middle school students. Students at different stages may participate in and react to the proposed guided inquiry approaches differently. Whether elementary school or high school students' scientific literacy will be influenced by the two instructional approaches requires further investigation. Gathering information on these issues through further studies can help to obtain a thorough understanding of the use of virtual labs in science classroom contexts, and thereafter an inquiry approach can be designed to enhance scientific literacy in a broader context.

Appendix 1: The Scientific literacy test related to Buoyancy

1. Eva conducted experiments to learn about Buoyancy in the laboratory. The experimental setup is illustrated in Fig. 5. Firstly, Eva weighed the stone by hanging it on a spring scale. The volume of the stone is 1200 cm^3 . It weighs 1500gw . Then, she lowered the stone into the water as shown in Step 2. In the third step, she recorded the weight of

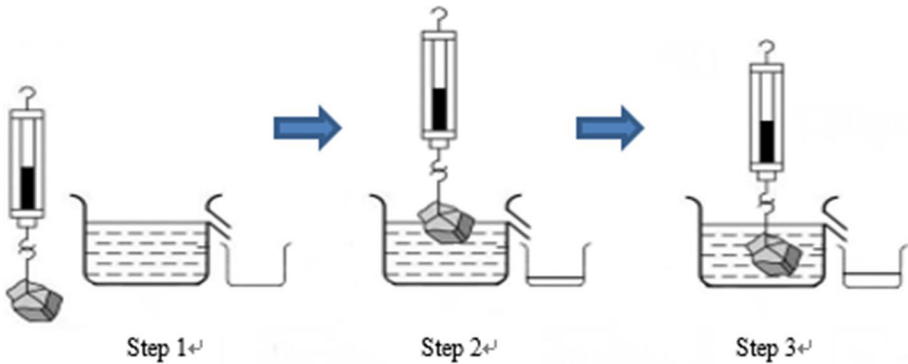


Fig. 5 The experimental setup and steps

the stone in the water and the volume of water that the stone displaced on the following sheet.

	Stone volume (cm ³)		Value on the spring scale (gw)	Weight of water displaced (gw)	Weight of the stone in water (gw)	Buoyancy force (gw)
	Under water	Above water				
Step1	0	1200	1500	0	Not in water	0
Step2	600	600	900	600	900	600
Step3	1200	0	300	1200	300	1200

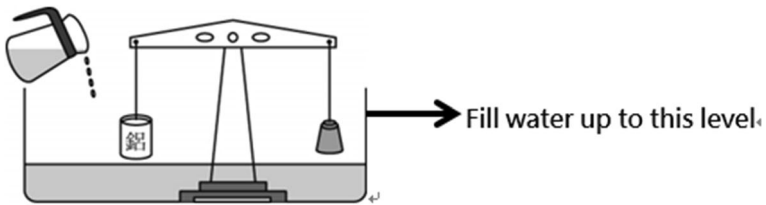
- 1-1 Observe the data from step 1 to step 3; what values do you find the same as buoyancy force? Why?
- 1-2 Observe the data from step 1 to step 3; explain how the weight of the stone in the water changed. Why?
- 1-3 When the volume of the stone in the water is 500cm³, what would be the weight measured by the spring scale? Predict and explain your answer.
(A) 1200gw (B) 1000gw (C) 600gw (D) 500gw

Eva wanted to know if the buoyancy force would be the same in different liquids. Four experiments were conducted by placing two different stones in water and alcohol and the following results were obtained:

	Exp.1	Exp.2	Exp.3	Exp.4
Liquid	Water	Water	Ethanol	Ethanol
Liquid density (g/cm ³)	1	1	0.8	0.8
Volume of stone (cm ³)	1200	800	1200	800
Volume of stone in water (cm ³)	1200	800	1200	800
weight of stone in water (gw)	300	200	540	360

- 1-4 Comparing the data of Exp.1 and Exp.2, what values of variables were changed?
weight of stone (B) liquid density (C) volume of stone in water (D) volume of stone

1-5 Eva observed that the buoyancy force changes when the same stone was immersed in different liquids. Could you support her observation by comparing data in her experimental results? Which experiments and variables would you choose to compare? Why?



Eva conducted experiments to learn about Buoyancy in the laboratory. The experimental setup is illustrated as above. She put an equal-arm balance in a water tank. She then hung an aluminum block on the left end of the arm and a balancing weight on the right end. The volume of the aluminum block is larger than the balancing weight. Then, she filled the tank with water. Please answer the following questions.

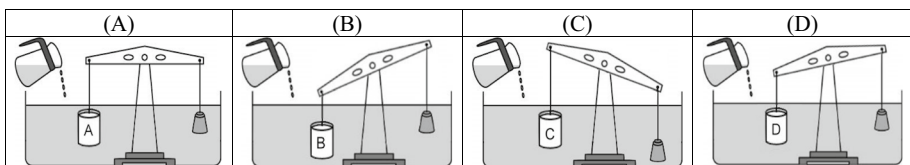
2-1 Which side do you think the arm would tip to? Why?

2-2 What would Eva's research question be for conducting this experiment?

2-3 Eva found some metals with the same mass as weight, but with different volumes. Since the labels had fallen off, Eva conducted the experiment using the device described above. She replaced the aluminum block with the metals A, B, C, and D, repeated the experiments and got the results shown in figures (A) to (D). Please identify the materials of blocks A to D and give explanations.

(The balancing weight is made of Iron. Its volume illustrated in the figures is not the actual size.)

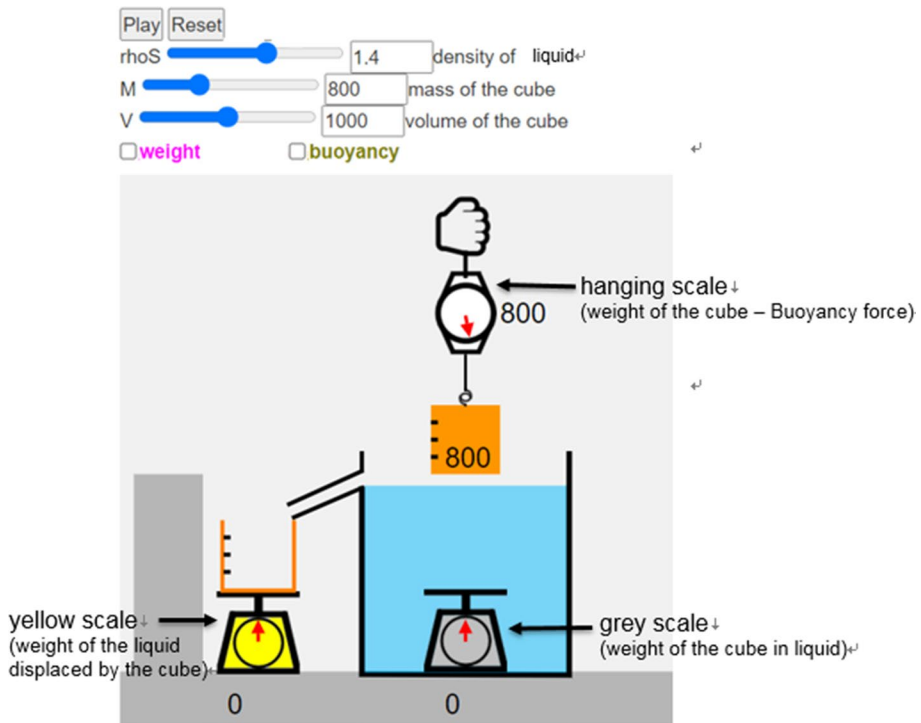
Material	Density (g/cm ³)
Platinum	21.5
Silver	10.5
Iron	7.9
Aluminum	2.7



- What is block A made of? Why?
- What is block B made of? Why?
- What is block C made of? Why?
- What is block D made of? Why?

Appendix 2: The critique worksheet

The physics virtual lab demonstrates the buoyancy in a liquid tank. You can change values of variables in the virtual lab to observe the process and conduct experiments about Buoyancy.



Datasheet

Variable	Description	Value	Unit
t	Time	0	second
rhoS	Density of the liquid	1.4	g/cm^3
M	Mass of the cube	800	g
V	Volume of the cube	1000	cm^3
rhoA	Density of the cube	0.8	g/cm^3
Vu	Volume of the cube under liquid	0	cm^3
Vs	Volume of liquid displaced by the cube	0	cm^3

Variable	Description	Value	Unit
scale1	Weight of the cube in liquid (grey scale)	0	gw
scale2	Weight of liquid displaced by the cube (yellow scale)	0	gw
scale3	Weight of the cube – Buoyancy force (hanging scale)	0	gw

Scenario 1

Ann wants to investigate how much Buoyancy the cube receives in liquid. She knows that Buoyancy force is the weight of the cube decreased in liquid. She would like to investigate “if the cube is easier to sink when it gets heavier.” Here are the experiments she designed:

Exp. 1	Exp. 2	Exp. 3	Exp.4	Exp. 5
rhoS= 1 (g/cm ³)	rhoS=0.8 (g/cm ³)	rhoS= 1 (g/cm ³)	rhoS=0.8 (g/cm ³)	rhoS= 1 (g/cm ³)
V = 1000 (cm ³)	V = 1000 (cm ³)	V = 1000 (cm ³)	V = 1000 (cm ³)	V = 1000 (cm ³)
M = 500 (g)	M = 600 (g)	M = 1000 (g)	M = 1000 (g)	M = 1200 (g)

1. Do you think Ann can examine her research question by the experiments she designed above? If not, how should she redesign them?
2. Data obtained from Ann’s experiments are listed below:

	rhoS liquid density	M cube’s mass	V cube’s volume	rhoA cube’s density	Vu cube’s volume under liquid	Vs liquid’s volume displaced by cube	Scale 1 grey scale	Scale 2 yellow scale	Scale 3 hanging scale
Exp. 1	1.00	500.00	1000.00	0.50	500.00	500.00	0.00	500.00	0.00
Exp. 2	0.80	600.00	1000.00	0.60	600.00	600.00	0.00	600.00	0.00
Exp. 3	1.00	1000.00	1000.00	1.00	1000.00	1000.00	0.00	1000.00	0.00
Exp. 4	0.80	1000.00	1000.00	1.00	1000.00	1000.00	0.00	800.00	200.00
Exp. 5	1.00	1200.00	1000.00	1.20	1000.00	1000.00	0.00	1000.00	200.00

2-1 Ann chooses Exp.3 and Exp.4 to examine her research question. Do you think her choice is right? Why?

2-2 If Ann chooses Exp.1 and Exp.2 and makes her conclusion: “the heavier the cube is, the more easily the cube sinks,” do you think her choice of experiments and conclusion are correct? Why? If she is wrong, what should she do?

2-3 In Exp. 1, Exp. 3, and Exp. 5, when the volume of the cube is fixed but the mass of the cube becomes heavier, which variable would change correspondingly?

Scenario 2

Hsian-Hsian would like to investigate how large a buoyancy force the cube receives in liquids. He knows that buoyancy force that an object receives equals the weight that it reduces after it submerges in liquid. He wants to know if Buoyancy force equals the weight of the liquid displaced by the cube. To test his thought, he designed the following experiments:

Exp. 1	Exp. 2	Exp. 3	Exp.4	Exp. 5
$\rho = 1 \text{ (g/cm}^3\text{)}$ $V = 1000 \text{ (cm}^3\text{)}$ $M = 600 \text{ (g)}$	$\rho = 1 \text{ (g/cm}^3\text{)}$ $V = 1000 \text{ (cm}^3\text{)}$ $M = 800 \text{ (g)}$	$\rho = 1 \text{ (g/cm}^3\text{)}$ $V = 1000 \text{ (cm}^3\text{)}$ $M = 1000 \text{ (g)}$	$\rho = 1 \text{ (g/cm}^3\text{)}$ $V = 1000 \text{ (cm}^3\text{)}$ $M = 1200 \text{ (g)}$	$\rho = 1.4 \text{ (g/cm}^3\text{)}$ $V = 1400 \text{ (cm}^3\text{)}$ $M = 1400 \text{ (g)}$

3. Data obtained from Hsian-Hsian’s experiments are listed below:

	ρ liquid den- sity	M cube’s mass	V cube’s volume	ρ_A cube’s den- sity	V_u cube’s volume under liquid	V_s liquid’s volume dis- placed by cube	Scale 1 grey scale	Scale 2 yellow scale	Scale 3 hang- ing scale	B buoy- ancy force
Exp. 1	1.00	600.00	1000.00	0.60	600.00	600.00	0.00	600.00	0.00	600
Exp. 2	1.00	800.00	1000.00	0.80	800.00	800.00	0.00	800.00	0.00	800
Exp. 3	1.00	1000.00	1000.00	1.00	1000.00	1000.00	0.00	1000.00	0.00	1000
Exp. 4	1.00	1200.00	1000.00	1.20	1000.00	1000.00	0.00	800.00	200.00	1000
Exp. 5	1.40	1400.00	1400.00	1.00	1000.00	1000.00	0.00	1400.00	0.00	1400

3-1 What variables should Hsian-Hsian choose to compare to test his hypothesis?

3-2 According to the results above, Hsian-Hsian claims that “I hypothesized that the buoyancy force equals the weight of the liquid displaced by the object. The experiment results support my hypothesis. So, I will not change my hypothesis. This is because the weight of the liquid displaced by the object equals the volume of the object under the liquid. The volume of the object under the liquid equals the buoyancy force received by the object. So the weight of the liquid displaced by the object equals the volume of the object under the fluid.”

Is this claim reasonable? If not, how should you revise the claim?

3-3 In fact, Hsian-Hsian’s result cannot support his hypothesis. Can you find out which experiment can not support his conclusion? Why?

Funding This work was supported by Ministry of Science and Technology, Taiwan, under Grants MOST106-2511-S-008-012-MY3, MOST107-2511-H-008-003-MY3, MOST107-2811-H-008-006, and MOST108-2628-H-003-001-MY3. It was also financially supported by the Institute for Research Excellence in Learning Sciences of National Taiwan Normal University (NTNU) from the Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan.

Data availability The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest All authors of this manuscript declare that we have no conflict of interest.

References

- Abdulwahed, M., & Nagy, Z. K. (2011). The TriLab, a novel ICT based triple access mode laboratory education model. *Computers & Education*, *56*(1), 262–274.
- Adriyawati, U., Rahmawati, Y., & Mardiah, A. (2020). STEAM-project-based learning Integration to Improve elementary school students' scientific literacy on alternative energy learning. *Universal Journal of Educational Research*, *8*(5), 1863–1873.
- Akaygun, S., & Adadan, E. (2019). Revisiting the understanding of redox reactions through critiquing animations in variance. In *Research and practice in chemistry education* (pp. 7–29). Springer
- Ardianto, D., & Rubini, B. (2016). Comparison of students' scientific literacy in integrated science learning through model of guided discovery and problem based learning. *Jurnal Pendidikan IPA Indonesia*, *5*(1), 31–37.
- Bouta, H., & Paraskeva, F. (2013). The cognitive apprenticeship theory for the teaching of mathematics in an online 3D virtual environment. *International Journal of Mathematical Education in Science and Technology*, *44*(2), 159–178.
- Cavagnetto, A. R. (2010). Argument to foster scientific literacy: A review of argument interventions in K–12 science contexts. *Review of Educational Research*, *80*(3), 336–371.
- Chang, C. J., Liu, C. C., Wen, C. T., Tseng, L. W., Chang, H. Y., Chang, M. H., et al. (2020). The impact of light-weight inquiry with computer simulations on science learning in classrooms. *Computers & Education*, *146*, 103770.
- Chang, H. Y. (2017). How to augment the learning impact of computer simulations? The designs and effects of interactivity and scaffolding. *Interactive Learning Environments*, *25*(8), 1083–1097.
- Chang, H. Y., & Chang, H. C. (2013). Scaffolding students' online critiquing of expert-and peer-generated molecular models of chemical reactions. *International Journal of Science Education*, *35*(12), 2028–2056.
- Chang, H.-Y., & Linn, M. C. (2013). Scaffolding learning from molecular visualizations. *Journal of Research in Science Teaching*, *50*(7), 858–886.
- Chao, J., Chiu, J. L., DeJaegher, C. J., & Pan, E. A. (2016). Sensor-augmented virtual labs: Using physical interactions with science simulations to promote understanding of gas behavior. *Journal of Science Education and Technology*, *25*(1), 16–33.
- Charney, J., Hmelo-Silver, C. E., Sofer, W., Neigeborn, L., Coletta, S., & Nemeroff, M. (2007). Cognitive apprenticeship in science through immersion in laboratory practices. *International Journal of Science Education*, *29*(2), 195–213.
- Collins, A. (2006). Cognitive apprenticeship. *The Cambridge Handbook of The Learning Sciences*, R. Keith Sawyer.
- Crujeiras-Pérez, B., & Jiménez-Aleixandre, M. P. (2017). High school students' engagement in planning investigations: Findings from a longitudinal study in Spain. *Chemistry Education Research and Practice*, *18*(1), 99–112.
- Dennen, V. P. (2000). Task structuring for on-line problem based learning: A case study. *Journal of Educational Technology & Society*, *3*(3), 329–336.
- Efstathiou, C., Hovardas, T., Xenofontos, N. A., Zacharia, Z. C., deJong, T., Anjewierden, A., et al. (2018). Providing guidance in virtual lab experimentation: The case of an experiment design tool. *Education Technology Research and Development*, *66*(3), 767–791.
- Ginns, P., Hu, F. T., Byrne, E., & Bobis, J. (2016). Learning by tracing worked examples. *Applied Cognitive Psychology*, *30*(2), 160–169.
- Henderson, J. B., MacPherson, A., Osborne, J., & Wild, A. (2015). Beyond construction: Five arguments for the role and value of critique in learning science. *International Journal of Science Education*, *37*(10), 1668–1697.

- Heradio, R., De La Torre, L., Galan, D., Cabrerizo, F. J., Herrera-Viedma, E., & Dormido, S. (2016). Virtual and remote labs in education: A bibliometric analysis. *Computers & Education*, *98*, 14–38.
- Hmelo, C., & Day, R. (1999). Contextualized questioning to scaffold learning from simulations. *Computers & Education*, *32*(2), 151–164.
- Hofstein, A., & Mamlok-Naaman, R. (2007). The laboratory in science education: The state of the art. *Chemistry Education Research and Practice*, *8*(2), 105–107.
- Ismail, I., Permanasari, A., & Setiawan, W. (2016). Stem virtual lab: An alternative practical media to enhance student's scientific literacy. *Jurnal Pendidikan IPA Indonesia*, *5*(2), 239–246.
- Jannati, E. D., Setiawan, A., Siahaan, P., & Rochman, C. (2018). Virtual laboratory learning media development to improve science literacy skills of mechanical engineering students on basic physics concept of material measurement. *Journal of Physics: Conference Series*, *1013*(1), 012061.
- Järvelä, S. (1995). The cognitive apprenticeship model in a technologically rich learning environment: Interpreting the learning interaction. *Learning and Instruction*, *5*(3), 237–259.
- Jin, G., & Bierma, T. (2013). STEM for non-STEM majors: Enhancing science literacy in large classes. *Journal of College Science Teaching*, *42*(6), 20–26.
- Karamustafaoglu, S. (2011). Improving the science process skills ability of science student teachers using I diagrams. *International Journal of Physics & Chemistry Education*, *3*(1), 26–38.
- Kusumastuti, F. A., Rombot, O., & Ariesta, F. W. (2019). The effect of stem integration on primary school students' scientific literacy. *International Journal of Scientific and Technology Research*, *8*(12), 1551–1553.
- Labouta, H. I., Kenny, N. A., Li, R., Anikovskiy, M., Reid, L., & Cramb, D. T. (2018). Learning science by doing science: An authentic science process-learning model in postsecondary education. *International Journal of Science Education*, *40*(12), 1476–1492.
- Lai, K., Cabrera, J., Vitale, J. M., Madhok, J., Tinker, R., & Linn, M. C. (2016). Measuring graph comprehension, critique, and construction in science. *Journal of Science Education and Technology*, *25*(4), 665–681.
- Larkins, D. B., Moore, J. C., Rubbo, L. J., & Covington, L. R. (2013, March). Application of the cognitive apprenticeship framework to a middle school robotics camp. In *Proceeding of the 44th ACM technical symposium on Computer science education* (pp. 89–94).
- Lefkos, I., Psillos, D., & Hatzikraniotis, E. (2011). Designing experiments on thermal interactions by secondary-school students in a simulated laboratory environment. *Research in Science & Technological Education*, *29*(2), 189–204.
- Li, H., Gobert, J., & Dickler, R. (2019, June). Scaffolding during Science Inquiry. In *Proceedings of the sixth (2019) ACM conference on learning@ scale* (pp. 1–10).
- Liu, T. C. (2005). Web-based cognitive apprenticeship model for improving pre-service teachers' performances and attitudes towards instructional planning: Design and field experiment. *Journal of Educational Technology & Society*, *8*(2), 136–149.
- Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research*, *28*(4), 587–604.
- Martínez, G., Naranjo, F. L., Perez, A. L., Suero, M. I., & Pardo, P. J. (2011). Comparative study of the effectiveness of three learning environments: Hyper-realistic virtual simulations, traditional schematic simulations and traditional laboratory. *Physical Review Special Topics-Physics Education Research*, *7*(2), 020111.
- Matuk, C., Zhang, J., Uk, I., & Linn, M. C. (2019). Qualitative graphing in an authentic inquiry context: How construction and critique help middle school students to reason about cancer. *Journal of Research in Science Teaching*, *56*(7), 905–936. <https://doi.org/10.1002/tea.21533>
- McElhane, K. W., Chang, H.-Y., Chiu, J. L., & Linn, M. C. (2015). Evidence for effective uses of dynamic visualizations in science curriculum materials. *Studies in Science Education*, *51*(1), 49–85.
- Moore, E. B., Smith, T. L., & Randall, E. (2016). Exploring the relationship between implicit scaffolding and inclusive design in interactive science simulations. In *International conference on universal access in human-computer interaction* (pp. 112–123). Springer
- Mørch, A. I., Engeness, I., Cheng, V. C., Cheung, W. K., & Wong, K. C. (2017). EssayCritic: Writing to learn with a knowledge-based design critiquing system. *Journal of Educational Technology & Society*, *20*(2), 213–223.
- Namdar, B., & Kucuk, M. (2018). Preservice science teachers' practices of critiquing and revising 5E lesson plans. *Journal of Science Teacher Education*, *29*(6), 468–484.
- Nolen, S. B., & Koretsky, M. D. (2018). Affordances of virtual and physical laboratory projects for instructional design: Impacts on student engagement. *IEEE Transactions on Education*, *61*(3), 226–233.
- OECD. (2016). *PISA 2015 Assessment and analytical framework: Science, reading, mathematics and financial literacy*. OECD Publishing.

- Oriol, M. D., Tumulty, G., & Snyder, K. (2010). Cognitive apprenticeship as a framework for teaching online. *MERLOT Journal of Online Learning and Teaching*, 6(1), 210–217.
- Osborne, J. (2014). Teaching scientific practices: Meeting the challenge of change. *Journal of Science Teacher Education*, 25(2), 177–196.
- Parker, W. C., & Hess, D. (2001). Teaching with and for discussion. *Teaching and Teacher Education*, 17(3), 273–289.
- Quellmalz, E. S., Silbergitt, M. D., Buckley, B. C., Loveland, M. T., & Brenner, D. G. (2020). Simulations for supporting and assessing science literacy. In *Learning and performance assessment: concepts, methodologies, tools, and applications* (pp. 760–799). IGI Global.
- Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., Soloway, E., et al. (2004). A scaffolding design framework for software to support science inquiry. *The Journal of the Learning Sciences*, 13(3), 337–386.
- Renken, M. D., & Nunez, N. (2013). Computer simulations and clear observations do not guarantee conceptual understanding. *Learning and Instruction*, 23, 10–23.
- Roth, W. M., & Roychoudhury, A. (1993). The development of science process skills in authentic contexts. *Journal of Research in Science Teaching*, 30(2), 127–152.
- Sever, S., Yurumezoglu, K., & Oguz-Unver, A. (2010). Comparison teaching strategies of videotaped and demonstration experiments in inquiry-based science education. *Procedia-Social and Behavioral Sciences*, 2(2), 5619–5624.
- Swaak, J., & de Jong, T. (2001). Learner vs system control in using online support for simulation-based discovery learning. *Learning Environments Research*, 4(3), 217–241.
- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. *The Cambridge Handbook of Multimedia Learning*, 3(2), 19–30.
- Taramopoulos, A., & Psillos, D. (2017). Complex phenomena understanding in electricity through dynamically linked concrete and abstract representations. *Journal of Computer Assisted Learning*, 33(2), 151–163.
- Thoms, L. J., & Girwidz, R. (2017). Virtual and remote experiments for radiometric and photometric measurements. *European Journal of Physics*, 38(5), 055301.
- van Joolingen, W. R., de Jong, T., Lazonder, A. W., Savelsbergh, E. R., & Manlove, S. (2005). Co-Lab: Research and development of an online learning environment for collaborative scientific discovery learning. *Computers in Human Behavior*, 21(4), 671–688.
- Vergara, D., Rubio, M. P., & Lorenzo, M. (2017). New approach for the teaching of concrete compression tests in large groups of engineering students. *Journal of Professional Issues in Engineering Education and Practice*, 143(2), 05016009.
- Wen, C. T., Chang, C. J., Chang, M. H., Chiang, S. H. F., Liu, C. C., Hwang, F. K., & Tsai, C. C. (2018). The learning analytics of model-based learning facilitated by a problem-solving simulation game. *Instructional Science*, 46(6), 847–867.
- Wen, C. T., Liu, C. C., Chang, H. Y., Chang, C. J., Chang, M. H., Chiang, S. H. F., Yang, K. F., & Hwang, F. K. (2020). Students' guided inquiry with simulation and its relation to school science achievement and scientific literacy. *Computers & Education*, 149, 103830.
- Wilkie, K. J., & Ayalon, M. (2019). Learning through critiquing: Investigating students' responses to others' graphs of a real-life functional situation. In *Annual conference of the Mathematics Education Research Group of Australasia 2019* (pp. 763–770). Mathematics Education Research Group of Australasia.
- Wu, H., Krajcik, J., & Soloway, E. (2002). Promoting conceptual understanding of chemical representations: Students' use of a visualization tool in the classroom. *Journal of Research in Science Teaching*, 38, 821–842.
- Yore, L. D., & Treagust, D. F. (2006). Current realities and future possibilities: Language and science literacy—empowering research and informing instruction. *International Journal of Science Education*, 28(2–3), 291–314.
- Zacharias, Z. C., Olympiou, G., & Papaevripidou, M. (2008). Effects of experimenting with physical and virtual manipulatives on students' conceptual understanding in heat and temperature. *Journal of Research in Science Teaching*, 45(9), 1021–1035.