Investigating the effects of structured and guided inquiry on students' development of conceptual knowledge and inquiry abilities: a case study in Taiwan

Su-Chi Fang, Ying-Shao Hsu, Hsin-Yi Chang, Wen-Hua Chang, Hsin-Kai Wu & Chih-Ming Chen

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION, 2016
http://dx.doi.org/10.1080/09500693.2016.1220688

Routledge
Taylor & Francis Group

Investigating the effects of structured and guided inquiry on students' development of conceptual knowledge and inquiry abilities: a case study in Taiwan

Su-Chi Fang[a], Ying-Shao Hsu[a], Hsin-Yi Chang[b], Wen-Hua Chang[a], Hsin-Kai Wu[a] and Chih-Ming Chen[c]

[a]Graduate Institute of Science Education, National Taiwan Normal University, Taipei, Taiwan; [b]Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taipei, Taiwan; [c]Graduate Institute of Library, Information and Archival Studies, National Chengchi University, Taipei, Taiwan

ABSTRACT

In order to promote scientific inquiry in secondary schooling in Taiwan, the study developed a computer-based inquiry curriculum (including structured and guided inquiry units) and investigated how the curriculum influenced students' science learning. The curriculum was implemented in 5 junior secondary schools in the context of a weeklong summer science course with 117 students. We first used a multi-level assessment approach to evaluate the students' learning outcomes with the curriculum. Then, a path analysis approach was adopted for investigating at different assessment levels how the curriculum as a whole and how different types of inquiry units affected the students' development of conceptual understandings and inquiry abilities. The results showed that the curriculum was effective in enhancing the students' conceptual knowledge and inquiry abilities in the contexts of the six scientific topics. After the curriculum, they were able to construct interconnected scientific knowledge. The path diagrams suggested that, due to different instructional designs, the structured and guided inquiry units appeared to support the students' learning of the topics in different ways. More importantly, they demonstrated graphically how the learning of content knowledge and inquiry ability mutually influenced one another and were reciprocally developed in a computer-based inquiry learning environment.

ARTICLE HISTORY
Received 27 January 2016
Accepted 1 August 2016

KEYWORDS
Instructional design; inquiry ability; multi-level assessment; path analysis; web-based inquiry curriculum

Introduction

Inquiry-based science has the potential to promote students' conceptual understanding, engagement in science and scientific literacy (e.g. Crawford, 2007; Minner, Levy, & Century, 2010). Reform documents and curriculum standards in science education have advocated the notion of teaching through inquiry in the past decades (National Research Council [NRC], 2000, 2012; Tytler, 2007). Echoing this prevailing view in science

CONTACT Ying-Shao Hsu yshsu@ntnu.edu.tw Department of Earth Sciences, Graduate Institute of Science Education, National Taiwan Normal University, 88, Section 4 Ting-Chou road, Taipei 116, Taiwan

education, the recent educational reform in Taiwan sets the goal to 'activate' teaching in secondary schools, which emphasises the need to engage students in scientific inquiry activities.

To promote inquiry-based science in secondary schooling, this study proposed an instructional design model to align inquiry activities with curriculum standards, and used computer technology to support the design, development and implementation of inquiry units. For the purpose of aligning inquiry activities with curriculum standards, the curriculum standards were first unpacked (Krajcik, McNeill, & Reiser, 2008). Together with inquiry abilities, the expected learning performances and the learning activities were developed. Therefore, the enactments of the learning activities helped students to not only learn scientific concepts, but also to develop inquiry abilities. The use of computer technology has made it possible to engage students in authentic scientific inquiry within classrooms (de Jong, 2006; Linn & Eylon, 2011a; van Joolingen, de Jong, & Dimitrakopoulou, 2007). The up-to-the-minute information provided by web-based learning environments enables teachers to monitor students' learning progress and provide timely feedback (Slotta & Linn, 2009; Zhang, Hsu, Wang, & Ho, 2015). For these reasons, the proposed instructional design model suggested infusing technology, such as simulations and animations, into inquiry tasks and constructing these tasks in a web-based learning environment.

Previous studies have provided substantial evidence that compared with traditional teaching approaches, inquiry-based instruction was more effective in enhancing students' science learning (Geier et al., 2008; Minner et al., 2010; Schroeder, Scott, Tolson, Huang, & Lee, 2007). As to different levels of inquiry (NRC, 2000), structured inquiry was frequently adopted for supporting students developing scientific explanations about collected data while guided inquiry was employed for equipping students with experimental abilities(e.g. Banchi & Bell, 2008). These studies provided fruitful findings about how one particular level of inquiry helped students' science learning. However, there has been little research comparing and delving into the effects of different levels of inquiry influence students' development of conceptual knowledge (CK) and inquiry ability (IA) (Bunterm et al., 2014). As science inquiry in the classroom takes different forms, it is important to examine the impacts of different levels of inquiry so that we can provide teachers with useful suggestions regarding when (in what situation or for what purpose) to implement which level of inquiry to best support students' science learning in the classroom.

This study proposes an instructional design model, and demonstrates how this model helped develop two types of inquiry learning units based on curriculum standards: structured inquiry units (SIUs) and guided inquiry units (GIUs). We intended to contribute to the understanding of how different levels of inquiry-based instruction influenced students' science learning in terms of the interplay between students' development CK and IAs. Specifically, we asked two research questions:

(1) How does the computer-based inquiry curriculum as a whole (six learning units) influence the students' development of CK and IAs?
(2) How do the different types of inquiry unit affect the students' development of CK and IAs?

Literature review

In order to close the gap between inquiry-based science and the science curriculum in everyday classrooms, the study proposed an instructional design model that offers a systematic way to transform the national curriculum standards into learning materials. To further understand the effects of inquiry activities on science learning, we developed two types of inquiry units and multi-level assessments, and used path analysis to reveal the complex interrelationships between the development of CK and IA in a computer-based learning environment. In this section, we first articulate the rationale behind the proposed instructional design model. Then, a critical review of relevant research about different types of classroom inquiry and measurement issues of learning outcomes with inquiry-based curriculum is presented.

Designing web-based science inquiry learning units based on standards

The national- or state-level science standards specify what science contents and abilities ought to be learned in a particular grade. Although the standards provide explicit direction for science teaching, the way they are presented poses challenges to the design of teaching materials. The standards present science contents as core ideas – for example, photosynthesis or force. Without systematic methods, it is particularly demanding for teachers to transform those core ideas into teaching materials or inquiry activities. To assist the transformation, we adopted and modified the method proposed by Krajcik et al. (2008). In brief, the core ideas are unpacked, delineated, and then integrated with IAs to form inquiry activities. During the process of unpacking, the concepts subsumed in the core ideas are specified with reference to the teachers' content knowledge and textbooks. By incorporating cognitive or inquiry elements, learning goals are established to describe how students are expected to perform their understandings of the concepts. Thus, apart from a full grasp of scientific terminology and basic concepts, the designed learning activities emphasise the practice of IAs, such as identifying trends in the data for proposing a claim or reasoning based on evidence. In other words, these activities are designed for students to *perform* how and what was understood. A more detailed introduction of the instructional design model will be presented in the section of Instructional Design.

For the past decades, rapid advances in technology have supported innovations in science learning and instruction. Inquiry teaching and learning also benefit greatly from the use of technology (de Jong, 2006; Krajcik, Czerniak, & Czerniak, 2007). Web-based learning environments are computer-based tools using hyperlinks to interconnect information structurally. Instead of presenting various software tools individually, web-based learning environments help integrate these tools and present them in a structured way. Together with embedded scaffolds and teacher guidance, web-based inquiry environments offer alternative ways to engage students in authentic scientific inquiry within classrooms (van Joolingen et al., 2007). The Collaborative Web-based Inquiry Science Environment (CWISE) plays a crucial part in this study (http://cwise.nccu.edu.tw). This online platform was originally developed by the University of Berkeley (named WISE) (Linn & Eylon, 2011b; Slotta & Linn, 2009). Our research team translated it into Chinese and is planning to extend its function to cater for collaborative learning. This learning environment not only allows us to adopt a variety of media to present inquiry activities, but also helps record students' learning process in real time. By scanning immediate records, teachers can monitor students' learning during the lessons. More importantly, teachers can give timely feedback to students and adjust the ongoing teaching process. Put differently, CWISE empowers teachers to engage students in scientific inquiry and enhances learning supports during inquiry activities.

A multi-level assessment approach to examine students' science learning

'In investigating inquiry learning we should not only look at the process of inquiry, but also to its result' (van Joolingen et al., 2007, p. 115). Indeed, one of the challenges the innovative curriculum faces is selection of trustworthy assessments to measure students' learning outcomes (Liu, Lee, & Linn, 2010). Previous research has pointed out that project-produced assessments tended to bias student learning outcomes towards positive learning (Slavin, 2008). This is because when students interact with the designed inquiry activities that are particularly tied to certain learning goals, they are likely to demonstrate improved learning outcomes on the assessments closely related to the learning activities. Therefore, it is suggested that assessments that are more distant from the enacted curriculum and set in different contexts are also needed to provide the evidence of learning progress (Geier et al., 2008; Liu et al., 2010).

The study adopted a multi-level assessment approach (Ruiz-Primo, Shavelson, Hamilton, & Klein, 2002) to examine the effectiveness of the web-based inquiry curriculum. Specifically, the assessments included two dimensions. The first dimension classifies the assessments according to two different facets of learning outcomes: CK and IA. The other dimension categorises the assessments into different levels 'based on their proximity to the enacted curriculum' (Krathwohl, 2002, p. 369). This study adopted three levels of assessment (immediate, close and proximal) to evaluate the students' learning outcomes. Immediate assessments served as pre-determined 'checking points', so that teachers could decide whether to proceed further or to give students more time and feedback. The assessments at the close level (i.e. the unit tests in our case) could provide a more complete view of the students' learning outcomes regarding each specific topic. The proximal-level assessments concerned the issue of learning transfer. That is, this level of assessment was used as an indicator to inform the influence of the curriculum on the students' development of CK and IA in a more general manner.

Different types of classroom inquiry and their impacts on students' science learning

Researchers have classified inquiry-based learning into different levels concerning varied instructional conditions (Bell, Smetana, & Binns, 2005; Buck, Bretz, & Towns, 2008; NRC, 2000; Tekkumru-Kisa, Stein, & Schunn, 2015). For example, NRC (2000) classified inquiry into four types as open, guided, structured and recipe-based, from the end that provides the least direction to the other end with more direction. Buck et al. (2008) used the amount of guidance given to students to define five levels of inquiry-based learning, from lower to higher levels, as confirmation, structured, guided, open and authentic inquiry. Each level is characterised by different provisions of guidance in the six areas: problem/question, theory/background, procedure/design, result analysis, results communication and conclusion. Similarly, the SIUs and GIUs in the present study were differentiated by the amount and the type of guidance provided. Specifically, in the SIUs, precise guidance to complete the learning activities was provided. The students were directed to understand the particular content, and engaged in inquiry tasks within the context of the same content. Comparatively, the guidance given in the GIUs was less and more general. The learning goals for the GIUs were to encourage students to apply what they had learned about the knowledge to solve a problem, explain phenomena or make decisions regarding complex socioscientific issues.

The variations in inquiry can be seen as a continuum, and it is assumed that students should progress gradually from lower to higher levels (NRC, 2000). Although the objective is to support students' development of knowledge and abilities to conduct inquiries at the highest level, they cannot be expected to begin there. Students need practice in inquiry and need to progressively build up their IAs. Previous research has suggested that providing explicit procedures and instruction is more effective in facilitating inquiry learning because the provision of supports helps reduce cognitive load and alleviates students' frustration (Furtak, Seidel, Iverson, & Briggs, 2012; Hmelo-Silver, Duncan, & Chinn, 2007). To implement classroom inquiry, however, the critical issue of when to implement which level of inquiry to best support students' science learning in classrooms needs to be addressed. Therefore, we argue that it is significant to understand the effects of different levels of inquiry on students' learning of content knowledge and IAs so that useful suggestions about inquiry can be provided. In reviewing the relevant literature, we learned that little research has looked into how different levels of inquiry influence students' science learning. Bunterm et al. (2014) adopted a quasi-experimental approach to investigate how structured and guided inquiry affected students' content knowledge, science process skills, scientific attitudes and self-perceived stress. In their study, The 5E Learning Cycle Model was used to design inquiry activities, and the difference between the two inquiry conditions was that the teachers gave specific instructions or provided guidance in the Engagement, Exploration and Explanation phases. The results showed that the students in the guided-inquiry condition had greater improvement in content knowledge and processing skills, but the two conditions did not seem to affect the students' scientific attitudes and stress in a consistent way. Sadeh and Zion (2009, 2011) compared how open-inquiry and guided-inquiry influenced students' dynamic inquiry performance and attitudes towards their inquiry projects. The students who learned in an open-inquiry environment had higher levels of performance in 'changes during inquiry' and 'procedural understanding', and expressed greater satisfaction with their learning projects. The aim of the present study was to contribute to the field regarding the effects of different levels of inquiry instruction on students' science learning. Particularly, we examined how students' acquisition of CK and progressions in IA influenced one another in a computer-based inquiry environment.

Method

Participants and procedure

The research team ran a four-day summer science course in five public junior high schools in northern, middle and southern Taiwan. The science course was open to all ninth graders in the schools, and the students volunteered to participate in the course with their parents' consent. A total of 117 ninth-grade students (effective sample size) completed the 6 learning units and pre- and post-assessments. Each learning unit lasted for three to four hours. The students took the pretests of the unit assessments and proximal assessments about a week before the inquiry curriculum. The unit post-tests were conducted immediately after each unit, and the proximal post-tests were administered at the end of the summer course.

Four earth science teachers, two physics teachers and two biology teachers taught the learning units in the five schools. Before implementing the inquiry curriculum, all of the participating teachers attended the workshops (two and half days) held by the research team. The workshops introduced the inquiry curriculum, provided teaching guidance, and created opportunities for teachers to discuss and share their ideas about how to implement the units. During the science course, the teachers also received technical support from the research team.

The instructional design model

The study used an instructional design model (Figure 1) to develop inquiry learning units on different scientific topics. Four key steps were included in the design process.

Step 1: Aligning standards and competencies. In the first step, the standards related to the topic were selected and unpacked into constituent concepts. Next, according to the conceptual hierarchy and the designer's content knowledge, the constituent concepts were structured to form a content map.[1] In order to decide how students are expected to perform what they are learning, IAs and basic cognitive processes were incorporated to establish the learning goals. In this study, three experimenting abilities (identifying and choosing variables (CV), planning an experiment (PE), transforming and representing data (TR)) and four explanatory abilities (making a claim (MC), using evidence (UV), describing the reasoning process (DR), and evaluating explanations (EE)) were included in designing the inquiry tasks (Hsu, Chang, Fang, & Wu, 2015). Apart from IAs, we incorporated two basic cognitive processes (remembering and understanding) (Krathwohl, 2002) into the expected learning performances. These basic cognitive performances were included for ensuring students' understanding of new terminology and helping make connections to prior knowledge. Table 1 shows two examples of developing learning performance and corresponding tasks.

Step 2 & Step 3: Planning inquiry tasks & Infusing technology and visualisation tools into learning tasks. As mentioned earlier, depending on the amount of direction provided by the material/the teacher or the extent to which learner self-direction is given, classroom inquiry can be classified into different levels (Bell et al., 2005; Buck et al., 2008; NRC, 2000). When designing inquiry tasks, one needs to decide which level of classroom inquiry he/she intends to shape, because this implicates how the scaffolds will be designed and the amount that is needed in the learning tasks. After deciding the level of inquiry, one can proceed to select suitable technology such as simulation or visualisation tools to develop inquiry tasks, for the purposes of achieving the learning goals determined in the previous step.

Step 4: Evaluating learning performances. In the fourth step, the learning units are implemented and refined based on the implementing results.

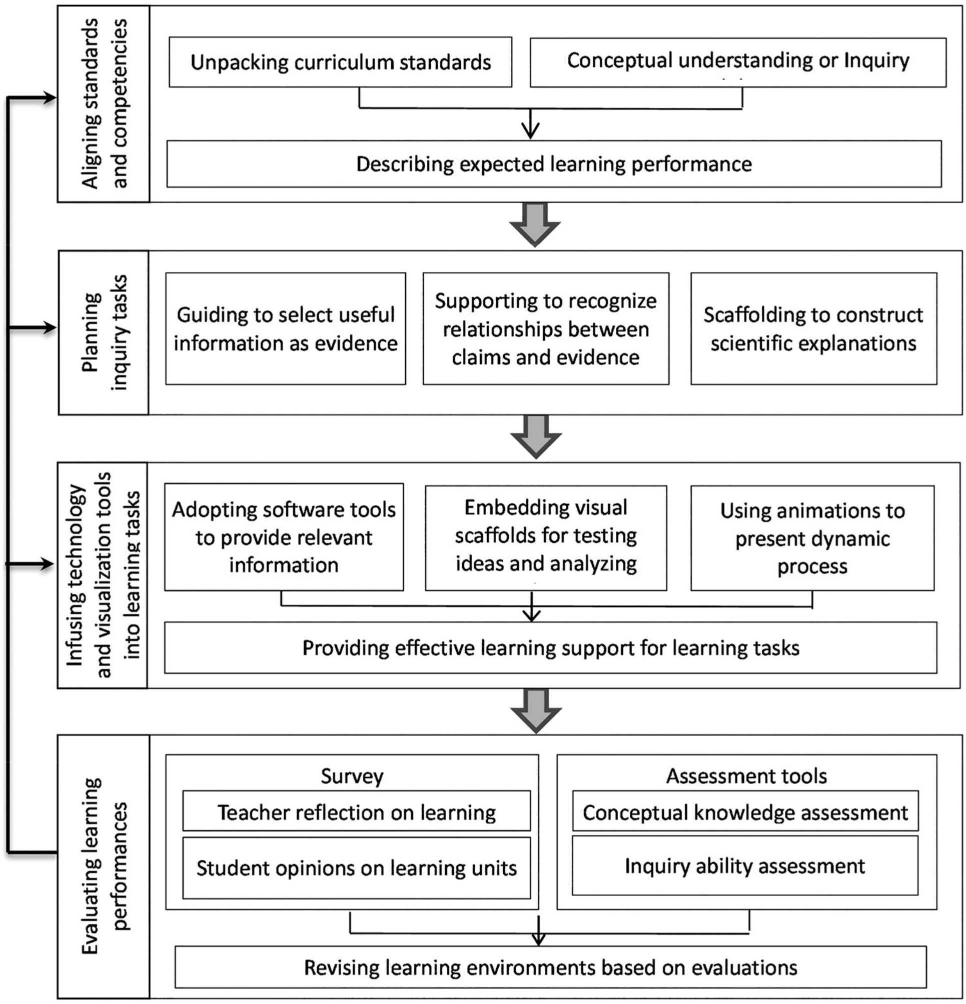


Figure 1. A design model of web-based inquiry learning curriculum.

The structured and Guided Inquiry Units

Focusing on the experimental and explanatory abilities (Hsu et al., 2015), our research team developed six inquiry learning units on CWISE. These units were classified into two inquiry types. The SIUs included three topics: Evolution, Buoyancy and Plate tectonics; the GIUs encompassed Genetics, Temperature and heat,[2] and Building a dam. The SIUs and GIUs in this study were different in two fundamental aspects: (1) the specificity and the range of the science contents involved and (2) the type of guidance designed in the material. Depending on the nature of the topic, the units each focused on different combinations of inquiry sub-abilities. Table 2 illustrates the inquiry sub-abilities involved in the learning units, and describes the design features of each unit.

Compared to the SIUs that focused on the particular concepts of a topic, the GIUs addressed the science contents across a range of topics (Kanter, 2010; Sherin, Edelson,

Table 1. Two examples of developing learning performance and its corresponding embedded assessment in the plate tectonics unit.

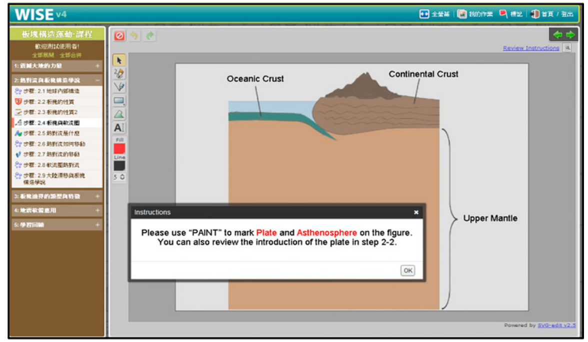
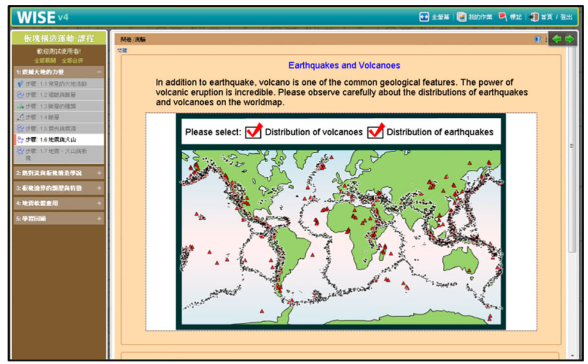
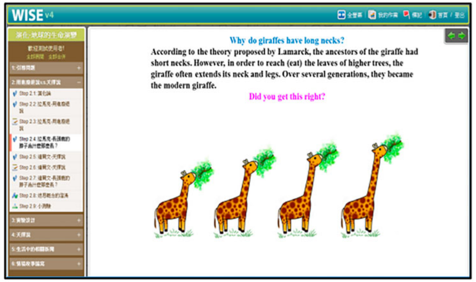
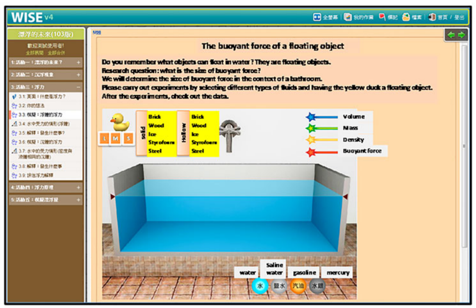
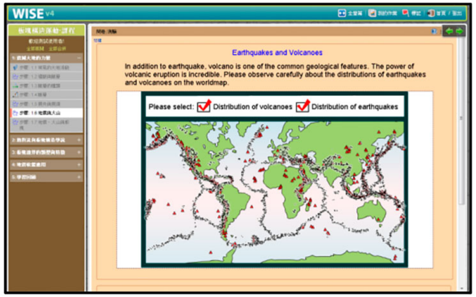
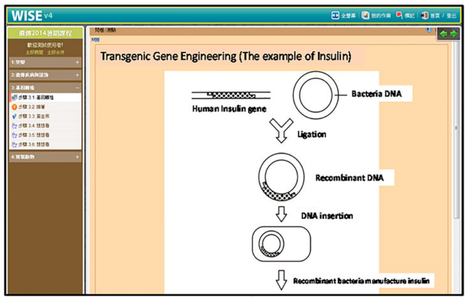
Unpacked contents	Cognitive process/IAs	Learning performances	Embedded assessment (conceptual item or inquiry item)
The definition of tectonic plate	Understanding	Be able to mark the positions of tectonic plates and the asthenosphere on a map of the earth's structure	
The relationships between seismic zone, volcanic belt and plate boundaries	Making a claim	Be able to deduce and make a claim regarding the relationships between tectonic plates, and volcanoes/ earthquakes from the given data	

Table 2. The inquiry sub-abilities involved and the design features of the six inquiry learning units.

Unit	Inquiry sub-abilities involved	Design features
Structured inquiry units		
Evolution 	Identifying and choosing variables (CV) Planning an experiment (PE), Making a claim (MC) Using evidence (UV) Evaluating explanations (EE)	Two theories about evolution: 'Natural selection' and 'Use and disuse', were introduced with the history of two biologists. The explaining power of the theories is further clarified through peer discussions, conceptual clarification and virtual experiments
Buoyancy 	Identifying and choosing variables (CV) Transforming and representing data (TR) Making a claim (MC) Using evidence (UV) Describing the reasoning process (DR) Evaluating explanations (EE)	The core concepts were introduced using virtual experiments. Students are guided to explain and summarise scientific principles with multiple representations. Then, students are required to apply the knowledge to an everyday life context: building a floating house

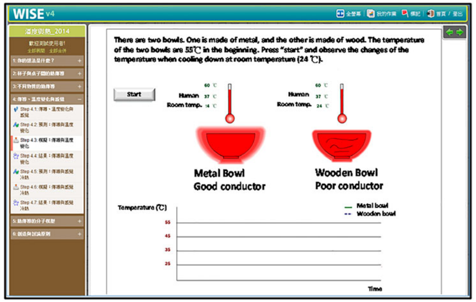
(*Continued*)

Table 2. Continued.

Unit	Inquiry sub-abilities involved	Design features
Plate Tectonics 	Making a claim (MC) Using evidence (UV) Describing the reasoning process (DR) Evaluating explanations (EE)	The core concepts were introduced with simulations. The software ‘Seismic Eruption’ provides real scientific records of volcanic eruptions and earthquakes for students to explore. Students were guided to observe the trends, analyse data and make claims based on the evidence
Guided inquiry units		
Genetics 	Transforming and representing data (TR) Making a claim (MC) Using evidence (UV) Describing the reasoning process (DR)	Relative concepts of gene mutation and heredity are introduced through the reading and discussion of related news and reports. The social-scientific issues related to these topics are also included in the unit

(*Continued*)

Table 2. Continued.

Unit		Inquiry sub-abilities involved	Design features
Temperature & heat		Identifying and choosing variables (CV) Planning an experiment (PE) Making a claim (MC) Using evidence (UV) Describing the reasoning process (DR) Evaluating explanations (EE)	Relative concepts are connected to everyday activities using simulations. Students are required to explain the particulate nature of matter with the support from visualisation tools. They are guided to integrate and reflect what they have learned for a deep conceptual understanding
Building a dam		Making a claim (MC) Using evidence (UV) Describing the reasoning process (DR) Evaluating explanations (EE)	'Reservoirs in Taiwan' provides information critical to building a dam. The dam software serves as a data resource providing necessary information for six locations. In the investigation process, students are guided to collect data, using evidence to support claims, and providing scientific explanations for their decisions

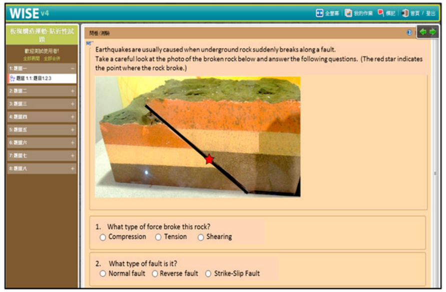
& Brown, 2004). The SIUs provided relevant data and asked learners to analyse them. They also provided explicit instruction for using evidence to formulate explanations and connect them to scientific knowledge. One SIU, plate tectonics, had particular content focuses, including the definition of a tectonic plate, the driving force of plate motion, the relationships between seismic zone, volcanic belt and plate boundaries, and three types of plate boundary (convergent, divergent and transform), etc. In addition to a basic introduction of these scientific concepts, several flash animations were used to support the students' reasoning of how the plates were forced, moved, and thus resulted in different geological phenomena. The software package 'Seismic Eruption' was also used as a data resource for exploring the relationships among seismic zones, volcanic belts and plate boundaries. Instead of giving explicit instructions, the GIUs directed and guided the learners to collect certain data, formulate explanations from evidence and connect explanations to scientific knowledge. 'The Dam' is one example. Learning this unit involved a broader range of concepts related to geology, topography, river systems, ecology and so on. The main activity was to select a dam location among six options based on a series of scientific investigations. The 'Reservoirs in Taiwan software' provided information and related knowledge critical to the construction of a dam. The dam software served as a data resource providing the necessary information for the six candidate locations. In the investigation process, the students were guided to collect relevant data, using evidence to support their claims, and providing scientific explanations for their decisions.

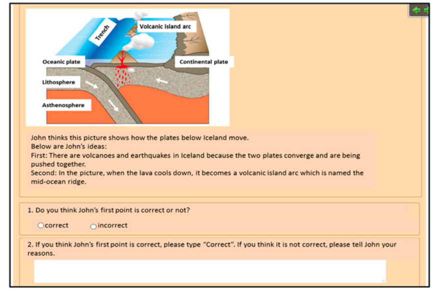
These 6 learning units were piloted in 2 secondary schools with a total of 68 eighth-grade students. All of them were refined based on the results of the pilot study.

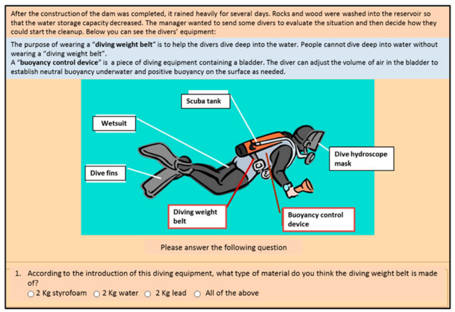
Research instruments

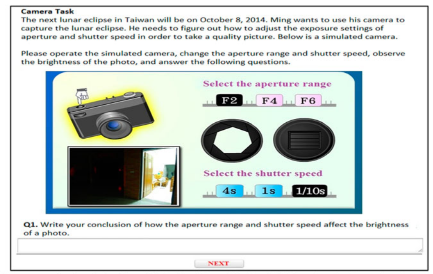
As foreshadowed in the literature review, the study designed assessments at the three levels: immediate, close and proximal (Ruiz-Primo et al., 2002). The development of immediate assessments was in parallel with the design of the learning unit and embedded in these units. To assure the alignment between the unit test and the learning unit, the unit tests were also developed by the experienced teachers who designed the inquiry learning units. The assessments at the close level (i.e. the unit tests) were used to evaluate the students' understanding of the same set of CK and IA. However, the questions at this level were designed in scenarios that were different from the learning units. They were reviewed for validity by the science education researchers in the research team, piloted with the learning units, and revised before the study. The proximal-level tests for CK and IA were independent. They were designed considering the knowledge and IAs related to the curriculum, but the content and the topics presented in the assessments were different (Ruiz-Primo et al., 2002). The proximal CK assessment used different topics to test the contents of all of the six topics. The test items were designed by a different group of experienced teachers and reviewed for validity by science education researchers. The proximal IA assessment was developed and validated by the assessment expert team (Kuo, Wu, Jen, & Hsu, 2015; Wu, Wu, & Hsu, 2014). Table 3 provides sample conceptual and inquiry test items from the close and proximal-level assessments.

Table 3. Sample conceptual and inquiry items of the close-level assessments (i.e. unit tests) and proximal-level assessments.

<table>
<tr><th>Level</th><th>Conceptual item</th><th>Inquiry item</th></tr>
<tr><td>Close-level assessment</td><td>


</td><td>


</td></tr>
<tr><td>Proximal-level assessment</td><td>


</td><td>


The screenshot is originally published in this article: Establishing the criterion-related, construct, and content validities of a simulation-based assessment of inquiry abilities. *International Journal of Science Education*, 36(10), 1630-1650.
</td></tr>
</table>

Data analysis

Regarding scoring, all of the multiple choice questions were scored dichotomously, with a score of 1 or 0 given for correct or incorrect answers. Short-answer conceptual questions were scored two, one, or zero respectively for complete, partial and irrelevant answers. We developed scoring rubrics for short-answer inquiry items (Kuo et al., 2015). A score of two was given for a high quality response, whereas scores of one and zero were given for moderate and low quality answers. All of the short-answer items were scored by two independent raters. The interrater reliability for the raters was found to be Kappa = 0.85–0.93. Since some of the test results were not normally distributed, Wilcoxon Signed-rank tests were adopted to examine the pre-/post-test gains on the unit tests, proximal conceptual and IA assessments.

To further explore how the students developed CK and IA during the inquiry curriculum, we adopted multiple regressions to conduct path analyses (Foster, Barkus, & Yavorsky, 2006; Streiner, 2005) to examine the relationships between the results of the multi-level assessments. Path analysis is a useful statistical approach for examining and revealing relative effects among variables (Foster et al., 2006; Lomax & Schumacker, 2012). Fundamentally, path analysis allows one to determine the relative contribution of some variables in predicting other variables through direct or indirect effects. A path diagram is first constructed to provide a graphic representation of theoretically hypothesised relationships among variables. There are two types of variables in a path diagram: exogenous and endogenous. Exogenous variables have outgoing arrows but none pointing at them, while endogenous variables have arrows pointing at them. Sometimes they have both incoming and outgoing arrows showing that they can be both independent and dependent variables (Klem, 1995). These arrows indicate the direction of effect. Calculation of standardised regression coefficients (i.e. path coefficient, β) estimates the degree and significance of the hypothesised relationship between the variables, while the squared multiple correlation (R^2) reflects the proportion of variance in the dependent variables accounted for by the proposed model. Three types of effects are generated from the analysis. A direct effect suggests one variable influences another without mediating variables, and the magnitude of the direct effect is represented by standardised regression coefficients (β). An indirect effect means the effect of one variable on another is mediated by at least one other variable in the system, and the effect is calculated by the sum of the products of direct effects through mediating variables. Thirdly, the total effect indicates the sum of the direct and indirect effects (Inan & Lowther, 2010).

For example, in this study, a path diagram for IA demonstrated in the unit post-tests (see Figure 2) was built based on a hypothesis that students' prior CK and prior IA in the context of the topic (exogenous variables), and their CK and IA demonstrated during the units (endogenous variables) were the key factors influencing their IA performed on the post unit test (endogenous variable) directly and indirectly. After the diagram was formed, the relationships between the variables in the model were designated by path coefficients showing the degree and the significance of the effects.

Results

Table 4 shows the students' performances on the unit tests before and after the inquiry curriculum. The results of the Wilcoxon signed-rank tests revealed that the students made significant improvements on all of the six unit tests from the pretest to the post-test with

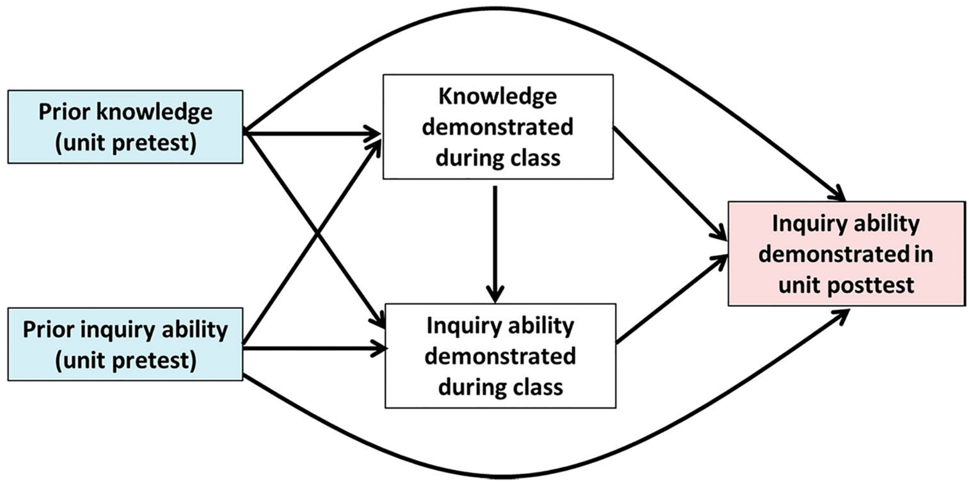


Figure 2. An example of the construction of a path diagram.

medium to large effect size. An independent analysis of the CK scores and IA scores also showed similar results (see Appendix 1). The findings indicated that the curriculum had promising effects on the students' CK and IA in the contexts of the six scientific topics.

Regarding the assessments at proximal level, the students had significant pre and post learning gains in the proximal CK assessments (Table 5), which implied that the six inquiry learning units supported the students' construction of robust and integrated conceptual understandings.

The proximal IA assessment result showed that there was no significant difference in the students' pre and post total inquiry scores. However, when the experimenting and explanatory ability scores were analysed separately, we found that the students made significant improvements in their experimenting abilities but not in their explanatory abilities (Table 6).

The above statistical comparisons of the pre- and post-assessments show that the students gained significant CK and IA in the context of the six units, and most of them were able to apply what they had learned to build integrated CK. Regarding these

Table 4. The six unit test results (close assessments).

					($N = 117$)		Wilcoxon signed-rank test		
		No. of items	Total score		*M*	*SD*	Z	*p*	*ES*[a]
Structured inquiry units	Evolution	35	62	Pre	21.24	8.46	6.70	<.001	0.62
				Post	26.44	8.97			
	Buoyancy	23	38	Pre	14.58	6.93	8.65	<.001	0.80
				Post	21.72	6.49			
	Plate tectonics	26	33	Pre	10.15	3.29	9.12	<.001	0.84
				Post	21.60	6.68			
Guided inquiry units	Genetics	23	37	Pre	19.48	5.35	7.61	<.001	0.70
				Post	24.75	5.19			
	Temperature and heat	19	28	Pre	16.88	4.78	3.96	<.001	0.37
				Post	18.34	4.81			
	Dam	30	45	Pre	17.64	8.73	8.14	<.001	0.75
				Post	26.09	7.65			

[a]Effect size $= Z/\sqrt{N}$.

Table 5. The proximal conceptual knowledge assessment results.

			$N = 117$		Wilcoxon signed-rank test		
No. of items	Total score		*M*	*SD*	*Z*	*p*	*ES*[a]
38	56	Pre	21.60	6.12	8.02	<.001	0.74
		Post	26.62	7.05			

[a]Effect size = $Z/\sqrt{N}$.

Table 6. The proximal inquiry ability assessment results.

						Wilcoxon signed-rank test		
	No. of items	Total score		*M*	*SD*	*Z*	*p*	*ES*[a]
Experimental abilities	16	26	Pre	14.53	5.01	2.05*	.041	0.19
			Post	15.36	5.15			
Explanatory abilities	17	33	Pre	14.76	5.26	0.49	.627	0.04
			Post	15.15	5.71			
Total inquiry ability score	33	59	Pre	29.28	9.20	1.55	.121	0.14
			Post	30.51	9.86			

[a]Effect size = $Z/\sqrt{N}$.

improvements, we used path analysis to explore (1) how the students' learning outcomes of the six inquiry units as a whole (close level) influenced the development of integrated CK (proximal level) and (2) how the students' learning performances during the SIUs and GIUs (immediate level) affected their learning outcomes of these two different types of inquiry unit (close level).

The path diagram in Figure 3 was built based on the assumption that the students' prior CK, prior IA, and the students' learning outcomes after the whole curriculum were the key factors influencing their performance on the proximal CK assessment. Among these factors, the students' prior CK and prior IA were measured by the proximal pretests, whereas the students' learning outcomes of the curriculum were measured by the unit post-tests. As shown in Figure 3, the students' prior CK and their IA demonstrated in the unit post-tests had significantly direct effects. It is worth noting that the students'

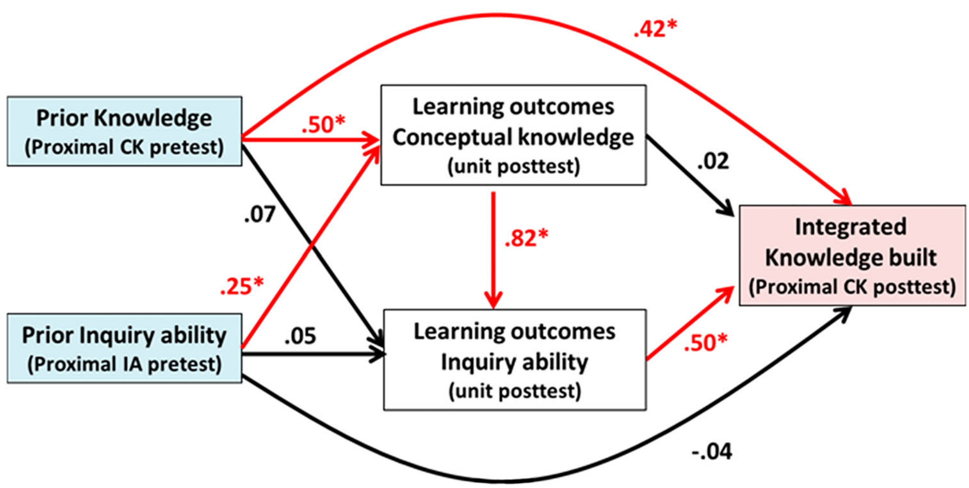


Figure 3. A path diagram for conceptual knowledge demonstrated in the proximal conceptual knowledge assessment.

prior CK, prior IA, and their CK demonstrated in the unit tests all had significant but indirect impacts through their IA demonstrated in the unit tests. These indirect impacts indicated that the students who had better prior CK and IA were likely to get higher scores in the conceptual items of the unit tests. Their rich conceptual understandings of the units in turn benefited their IA performances and were closely associated with their construction of integrated CK. Regarding the total significant effects (Appendix 2), the students' prior knowledge had the strongest effect on their proximal CK post-test (.62). The students' CK (.41) and IA (.50) demonstrated in the unit tests had medium effects, while their prior IA (.10) had the least effect.

Since the SIUs and GIUs were very different in terms of the contents and guidance designed in the learning materials, we conducted path analyses for these two different inquiry units. When determining the meaningful learning path to the learning outcomes of the six topics, we considered four key factors: the students' prior CK and prior IA in the contexts of the six topics, and their CK and IA demonstrated during the units. These factors were measured by the conceptual items and inquiry items in the unit pretests and in the embedded assessments accordingly. Figure 4 shows the analysing results.

When learning in the SIUs, the students' CK demonstrated during class had significantly direct effects on both their CK and IA performed in the unit post-tests. In addition, the other three variables, the students' prior CK, prior IA, and their IA demonstrated during class, all had indirect but significant effects through their knowledge demonstrated during class. This suggests that the students with higher prior CK and IA were likely to perform better during the structured units. Also, the students' advanced CK during class was closely related to their higher scores in the unit post-tests. The difference between the development of CK and IA in the SIUs was that the former was significantly

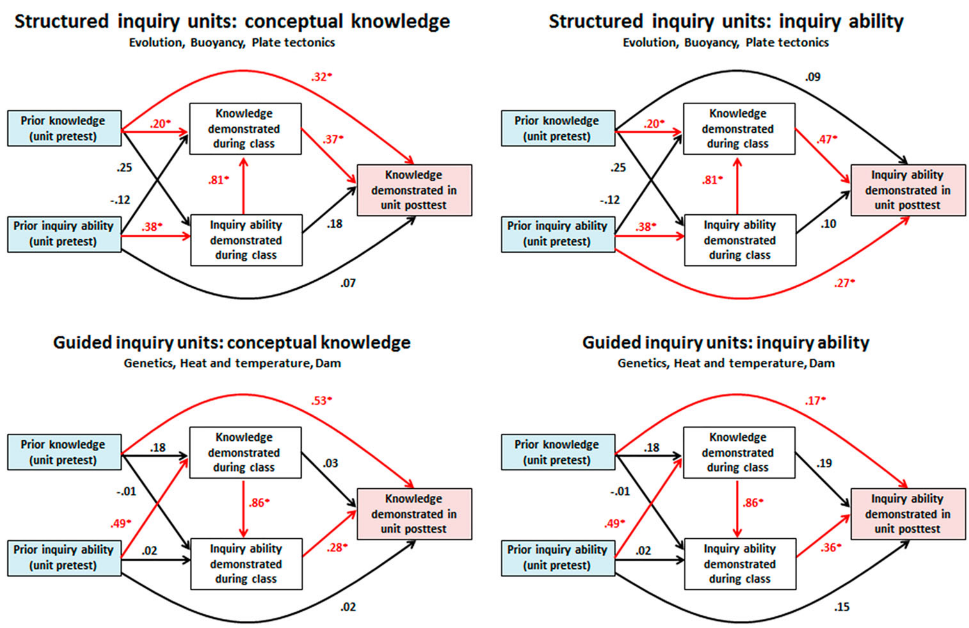


Figure 4. Path diagrams for conceptual knowledge and inquiry ability demonstrated in structured and GIU post-tests.

influenced by prior knowledge, while the latter was influenced by prior IA. Regarding the total effects (Appendix 3), the students' prior knowledge (.39) and the knowledge demonstrated during class (.37) had the strongest effects on their knowledge learning outcomes. The students' prior IA (.41) and their learning of the units (knowledge = .47, IA = .38) all had stronger effects on their inquiry learning outcomes.

Regarding the GIUs, the students' prior CK and their IA demonstrated during class had significantly direct effects on both their CK and IA performed in the unit post-tests. Unlike SIUs, only two variables (the students' prior IA and their CK) demonstrated during class intensified the effect of the IA demonstrated during class. These indirect effects implied that the students' prior IA was likely to support their knowledge construction, and then benefited their IA performed during the units. The better the IA the students exhibited during the GIUs, the higher the scores they gained in the unit post-tests. With respect to the total effects (Appendix 4), despite the students' prior CK (.53) having the strongest effects on their knowledge learning outcomes, what they had learned during class (knowledge = .24; IA = .28) made a significant contribution to their higher scores in the knowledge items. In contrast, the students' prior CK (.17) and prior IA (.15) had weaker effects on their inquiry performances in the unit post-test, but what they had learned during class (knowledge = .31; IA = .36) had the strongest impacts.

On the whole, the results suggested that in the SIUs, the knowledge the students constructed during class played a direct and critical role in facilitating their learning performance, while the IA they developed during class was rather auxiliary. By contrast, in the GIUs, what the students learned about IA during class was more significant in predicting their learning outcomes.

Discussion

This study demonstrates a way to develop an inquiry-based curriculum based on curriculum standards, and explored mutual influences between the students' development of CK and IA in a computer-based inquiry environment. The findings show that the curriculum helped construct integrated scientific knowledge, and significantly improved the students' CK and IA in the contexts of the six scientific topics. However, the students did not seem to be able to 'transfer' what they had learned about the IA to different settings. From the path analysis, we found that the students' construction of integrated knowledge was closely related to their prior CK and the IA developed during class. Moreover, owing to different instructional designs, the SIUs and GIUs appeared to support the students' learning of the topics in different ways.

The issue of 'learning transfer' in inquiry-based science learning

As previous research has pointed out, the assessments at a close level were tied to the learning units and were highly contextualised; they were rather sensitive to the students' learning progress (Ruiz-Primo et al., 2002) compared to those at a more distant level. The nature of close-level assessment explains the relatively higher effect sizes in the statistical comparison of pre and post unit tests in the study. To provide further evidence of students' learning progress, we used two independent assessments at the proximal level to examine the students' development of CK and IA. As the proximal CK assessment involved joined scientific contents, the students' significant improvements in this assessment indicated that the inquiry curriculum was effective in fostering robust and integrated conceptual learning. This in fact corresponds to previous research findings that inquiry-based instruction has great potential to assist students in building deep and interconnected CK (Fogleman, McNeill, & Krajcik, 2011; Hmelo-Silver et al., 2007; Linn & Eylon, 2011a). By means of path analysis, we took a further step to investigate in what way the integrated CK was influenced by the unit conceptual and inquiry learning outcomes. The results pointed out that the students' development of IA was indeed one critical variable affecting their construction of integrated CK. Therefore, we believed that situating content learning in inquiry activities can not only provide opportunities for fostering IAs, but can also help construct deep, meaningful and interconnected conceptual understandings (e.g. Geier et al., 2008; Minner et al., 2010).

Regarding the transfer of IA to new settings, the findings showed that although the students did not make significant progress in their explanatory abilities, they had significantly better experimenting abilities after completing the curriculum. Three units (heat and temperature, buoyancy, and evolution) engaged the students in practising experimenting abilities. One common feature of these units was that they all used visualisation tools to support the investigation process, such as identifying and choosing variables or planning experiments. For example, in the buoyancy unit, a simulation was developed for students to conduct virtual experiments. Students can manipulate relevant variables, such as the size of the object or the type of fluid, and use virtual probes to measure the properties of the selected object. By manipulating the simulation, analysing data and synthesising the results, students were expected to learn the key factors related to an object sinking or floating by conducting. After comprehending these fundamental ideas, they were required to apply what they had learned about buoyancy to an everyday context: how to build a floating house (Chang, Hsu, Wu, & Chen, 2014). We believe that the use of the visualisation tools in the study provided effective scaffolds for investigating process, which might have led to the significant improvement in the students' experimenting abilities (Linn, Bell, & Davis, 2004; Quintana et al., 2004).

Compared to experimenting abilities, performing explanatory abilities appears to be even more tightly connected to the scientific contents and theories. Performing explaining abilities involves the active search for and application of relevant CK to construct explanations (Eberbach & Crowley, 2009; Manz, 2012). Students are required to not only have a fully conceptual understanding of the scientific concepts, but also the ability to identify relevant evidence and make logical connections between the evidence and explanations (Kyza, Constantinou, & Spanoudis, 2011; Wu & Hsieh, 2006). Consequently, learning transfer of explanatory abilities to different contexts is indeed challenging for students, and this might be one reason explaining why the students in the study made significant progress in experimenting abilities but not in explanatory abilities. van Joolingen et al. (2007) pointed out that students may have better performance with support in the learning process (the first order effect); however, this does not necessarily lead to a better post-test score (the second order effect). Accordingly, the students' modest performance of explanatory abilities might indicate that they need more time to develop their abilities. We should not expect that sophisticated IAs can be developed only through a few lessons. Instead, it is essential to place inquiry-based instructional design at a curricular level (van Joolingen et al., 2007, p. 113) so that the students can be involved in inquiry activities for a longer period of time.

How the SIUs and GIUs influenced the development of CK and IAs

In the study, we found that the SIUs and GIUs appeared to support the students' learning in distinct ways. One obvious distinction was that the students' learning of the CK during class was pivotal for them to achieve successful learning of the SIUs, whereas the students' development of IA during class was crucial for the mastery of the GIUs. We speculated that these diverse effects sprang from the differences in the design of the two types of inquiry units. The SIUs and GIUs were different in terms of the specificity and the range of the science contents involved, and the amount of guidance designed in the material. The SIUs had particular content focuses and precise guidance for the inquiry activities. They were likely to draw the students' attention to making meaning of the particular CK. Thus the students with higher prior CK and prior IA were able to learn better CK during class, which in turn leveraged the learning outcomes of the SIUs. On the other hand, the practice of inquiry appeared to play an auxiliary but important role in assisting the students in sharpening and integrating their understanding of the scientific concepts. By contrast, the GIUs involved a broader range of content ideas and more general guidance for the inquiry activities. Hence they were likely to 'activate' the students' application of IA. In other words, the GIUs that featured relatively less emphasis on the learning of particular content ideas were inclined to draw students' attention to the inquiry practice itself. Therefore, the higher the prior IA, the better they were able to learn and apply the knowledge, and thus the better the IA they performed during the class, the higher the scores they gained in the unit post-tests.

In sum, one notable finding from the path analysis was that the significantly direct effects shown in the models in fact reflected different emphases in the design of the SIUs and GIUs. That is, the SIUs stressed a conceptual understanding of the particular concepts, whereas the GIUs highlighted the performances of IAs. Moreover, it should be noted that the students' learning of CK and IA during class mutually influenced one another, and these mutual effects also had significant impacts, directly or indirectly, on their science learning in a computer-based inquiry environment.

Limitations

Path analysis is a useful approach for analysing which variables exert effects on others. One should note that this approach cannot be used to establish or test causality. It is a technique for testing models, not for building them (Foster et al., 2006). Therefore, although the path models suggested in the study were based on theory, they were not the only ones or 'perfect' models because the analysis did not deal with the issue of model fit. What the analysis focused on was examining the hypothesised relationships between variables. Regarding the research design, the study did not compare two intervention groups using different inquiry instructions. Also, since a series of six units was taught consecutively, the data collected in regard to the SIUs or GIUs might have been affected by what the students had learned in previous units, and this might disturb the reported effects among variables in the study.

Conclusion

Instead of seeing science as merely a body of knowledge, recent reform in science education devotes attention to the notion of 'science-as-practice' that views science as a particular epistemic, social and cultural practice (Erduran & Dagher, 2014). To shift the focus towards 'practice', the American standards, A Framework for K-12 Science Education, suggested that science and engineering education should 'support the integration of such knowledge and abilities with the practices needed to engage in scientific inquiry and engineering design' (NRC, 2012, p. 2). In response to this shifted focus, the study offers a systematic way to link content knowledge and IAs in instructional design, and takes advantage of the web-based learning environment to enhance students' engagement in scientific practices rather than drilling isolated science facts or processing skills. The multi-level assessment and path analysis approaches used in the study offer an alternative way to look into the interplay between CK and IAs developed in a computer-based inquiry environment. We hope that, apart from supporting teachers' development of inquiry teaching materials, the study can also serve as a foundation empowering teachers to gradually and increasingly infuse inquiry elements into their everyday science classrooms. Future research can adopt a longitudinal approach to further dig into students' science learning during inquiry-based curricula, and continue contributing to the knowledge of the design challenges in devising inquiry instruction (Kanter, 2010; Meyer, AntinkMeyer, Nabb, Connell, & Avery, 2013).

Notes

1. A complete unpacking results and content map can be found in the book chapter: Developing technology-infused inquiry learning environment to promote science learning in Taiwan (Hsu et al., 2015).
2. This unit was originally designed by the WISE team (Chang & Linn, 2013).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Ministry of Science and Technology, Taiwan [grant number NSC-102-2811-S-003-003].

Notes on contributors

Su-Chi Fang is an assistant professor in Graduate Institute of Science Education, National Taiwan Normal University. She received her PhD in science education from University of Melbourne. Her dissertation is about teaching and learning a chemistry concept, the mole concept, in Australia and Taiwan. She taught science at secondary levels for eight years before becoming an educational researcher. Her research interests include classroom research, international comparison, inquiry teaching and learning, and practical work in science education.

Ying-Shao Hsu is a Professor of Graduate Institute of Science Education and the Department of Earth Sciences at the National Taiwan Normal University (NTNU). Currently, she is also the director of Graduate Institute of Science Education at present and Research Chair Professor of NTNU. She received her Ph.D. degree in 1997 from the Department of Curriculum and Instruction at the Iowa State University. Her expertise in research includes: technology infused learning, inquiry learning, metacognition, and earth science education. She has published journal articles in leading journals of science education such as Science Education, International Journal of Science Education, British Journal of Educational Technology, International Journal of Science and Mathematics Education, and Research in Science Education. She currently serves as an Editorial Board member of International Journal of Science and Mathematics Education, and as a reviewer for several SSCI journals. She is a recipient of Outstanding Research Award of the Ministry of Science and Technology, R.O.C. in 2012 and 2016.

Hsin-Yi Chang (Ph.D. from University of Michigan) is an associate professor at National Taiwan University of Science and Technology, Taiwan. Her research involves the use of simulation, visualization and augmented reality technology to support science learning. Her work has been published in Journal of Research in Science Teaching, Science Education, Computers & Education, and Instructional Science among others.

Wen-Hua Chang is Professor of the GISE, NTNU, Taiwan.

Hsin-Kai Wu is a Research Chair Professor in Graduate Institute of Science Education at the National Taiwan Normal University. She received her PhD in science education from the University of Michigan. Her research interests include science visualization, multiple representations, learning technology, and inquiry learning. She is an associate editor of International Journal of Science and Mathematics Education and serving as an editorial board member of Science and American Education Research Journal.

Chih-Ming Chen is a distinguished professor in the Graduate Institute of Library, Information and Archival Studies at National Chengchi University, Taipei, Taiwan. He received B.Sc. and M.Sc. degree from the Department of Industrial Education at National Taiwan Normal University in 1992 and 1997, and received Ph.D. degree from the Department of Electronic Engineering at National Taiwan University of Science and Technology in 2002. His research interests include digital library, e-learning, data mining, machine learning and intelligent agents on the web.

ORCiD

Hsin-Kai Wu http://orcid.org/0000-0003-0018-9969

References

Banchi, H., & Bell, R. (2008). The many levels of inquiry. *Science and Children*. http://search.proquest.com/docview/236901022?accountid=14228

Bell, R., Smetana, L., & Binns, I. (2005). Simplifying inquiry instruction. *The Science Teacher*, *72*(7), 30–33.

Buck, L. B., Bretz, S. L., & Towns, M. H. (2008). Characterizing the level of inquiry in the undergraduate laboratory. *Journal of College Science Teaching*, *38*(1), 52–58.

Bunterm, T., Lee, K., Ng Lan Kong, J., Srikoon, S., Vangpoomyai, P., Rattanavongsa, J., & Rachahoon, G. (2014). Do different levels of inquiry lead to different learning outcomes? A comparison between guided and structured inquiry. *International Journal of Science Education*, *36*(12), 1937–1959.

Chang, H.-Y., Hsu, Y.-S., Wu, H.-K., & Chen, C.-M. (2014). *Path analyses of how students develop conceptual knowledge and inquiry skills in a simulation-based inquiry environment*. Paper presented at the the 22nd International Conference on Computers in Education, Japan.

Chang, H.-Y., & Linn, M. C. (2013). Scaffolding learning from molecular visualizations. *Journal of Research in Science Teaching*, *50*(7), 858–886.

Crawford, B. A. (2007). Learning to teach science as inquiry in the rough and tumble of practice. *Journal of Research in Science Teaching*, *44*(4), 613–642.

Erduran, S., & Dagher, Z. R. (2014). Scientific Practices. In S. Erduran & Z. R. Dagher (Eds.), *Reconceptualizing the Nature of Science for Science Education: Scientific Knowledge, Practices and Other Family Categories* (pp. 67–89). Dordrecht: Springer.

Eberbach, C., & Crowley, K. (2009). From everyday to scientific observation: How children learn to observe the biologist's world. *Review of Educational Research*, *79*(1), 39–68.

Fogleman, J., McNeill, K. L., & Krajcik, J. (2011). Examining the effect of teachers' adaptations of a middle school science inquiry-oriented curriculum unit on student learning. *Journal of Research in Science Teaching*, *48*(2), 149–169.

Foster, J. J., Barkus, E., & Yavorsky, C. (2006). Path analysis. In J. J. Foster, E. Barkus, & C. Yavorsky (Eds.), *Understanding and using advanced statistics* (pp. 89–102). London: SAGE.

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research*, *82*(3), 300–329.

Geier, R., Blumenfeld, P. C., Marx, R. W., Krajcik, J. S., Fishman, B. J., 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.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99–107.

Hsu, Y.-S., Chang, H.-Y., Fang, S.-C., & Wu, H.-K. (2015). Developing technology-infused inquiry learning environment to promote science learning in Taiwan. In M. S. Khine (Ed.), *Science education in East Asia: Pedagogical innovations and best practices* (pp. 373–403). Cham, Switzerland: Springer.

Inan, F. A., & Lowther, D. L. (2010). Laptops in the K-12 classrooms: Exploring factors impacting instructional use. *Computers & Education*, *55*(3), 937–944.

de Jong, T. (2006). Computer simulations – technological advances in inquiry learning. *Science*, *312*, 532–533.

van Joolingen, W. R., de Jong, T., & Dimitrakopoulou, A. (2007). Issues in computer supported inquiry learning in science. *Journal of Computer Assisted Learning*, *23*(2), 111–119.

Kanter, D. E. (2010). Doing the project and learning the content: Designing project-based science curricula for meaningful understanding. *Science Education*, *94*(3), 525–551.

Klem, L. (1995). Path analysis. In L. G. Grimm, & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 65–98). Washington, DC: American Psychological Association.

Krajcik, J., Czerniak, C. M., & Czerniak, C. L. (2007). Using learning technology to support students in inquiry. In *Teaching science in elementary and middle school: A project-based approach* (3rd ed., pp. 163–192). New York, NY: Routledge.

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.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, *41*(4), 212–218.

Kuo, C.-Y., Wu, H.-K., Jen, T.-H., & Hsu, Y.-S. (2015). Development and validation of a multimedia-based assessment of scientific inquiry abilities. *International Journal of Science Education*, *37*(14), 2326–2357.

Kyza, E. A., Constantinou, C. P., & Spanoudis, G. (2011). Sixth graders' co-construction of explanations of a disturbance in an ecosystem: Exploring relationships between grouping, reflective scaffolding, and evidence-based explanations. *International Journal of Science Education*, *33*(18), 2489–2525.

Linn, M. C., Bell, P., & Davis, E. A. (2004). Specific design principles: Elaborating the scaffolded knowledge integration framework. In M. C. Linn, E. A. Davis, & P. Bell (Eds.), *Internet environments for science education* (pp. 315–339). Mahwah, NJ: Lawrence Erlbaum Associates.

Linn, M. C., & Eylon, B.-S. (2011a). Making visualizations valuable. In *Science learning and instruction: Taking advantage of technology to promote knowledge integration* (pp. 186–217). New York, NY: Routledge.

Linn, M. C., & Eylon, B.-S. (2011b). *Science learning and instruction: Taking advantage of technology to promote knowledge integration*. New York, NY: Routledge.

Liu, O. L., Lee, H.-S., & Linn, M. C. (2010). Multifaceted assessment of inquiry-based science learning. *Educational Assessment*, *15*(2), 69–86.
Lomax, R. G., & Schumacker, R. E. (2012). *A beginner's guide to structural equation modeling*. New York, NY: Routledge Academic.
Manz, E. (2012). Understanding the codevelopment of modeling practice and ecological knowledge. *Science Education*, *96*(6), 1071–1105.
Meyer, D., AntinkMeyer, A., Nabb, K., Connell, M., & Avery, L. (2013). A theoretical and empirical exploration of intrinsic problems in designing inquiry activities. *Research in Science Education*, *43*(1), 57–76.
Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction – what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching*, *47*(4), 474–496.
National Research Council. (2000). *Inquiry and the national science education standards: A guide for teaching and learning*. Washington, DC: National Academies Press.
National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.
Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., … Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences*, *13*(3), 337–386.
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.
Sadeh, I., & Zion, M. (2009). The development of dynamic inquiry performances within an open inquiry setting: A comparison to guided inquiry setting. *Journal of Research in Science Teaching*, *46*(10), 1137–1160.
Sadeh, I., & Zion, M. (2011). Which type of inquiry project do high school biology students prefer: Open or guided? *Research in Science Education*, *42*(5), 831–848.
Schroeder, C. M., Scott, T. P., Tolson, H., Huang, T.-Y., & Lee, Y.-H. (2007). A meta-analysis of national research: Effects of teaching strategies on student achievement in science in the United States. *Journal of Research in Science Teaching*, *44*(10), 1436–1460.
Sherin, B. L., Edelson, D., & Brown, M. (2004). On the content of task-structured curricula. In L. B. Flick, & N. G. Lederman (Eds.), *Scientific inquiry and nature of science: Implications for teaching, learning, and teacher education* (pp. 221–248). Dordrecht: Springer.
Slavin, R. E. (2008). Perspectives on evidence-based research in education – what works? Issues in synthesizing educational program evaluations. *Educational Researcher*, *37*(1), 5–14.
Slotta, J. D., & Linn, M. C. (2009). *WISE science: Web-based inquiry in the classroom*. New York, NY: Teachers College Press.
Streiner, D. L. (2005). Finding our way: An introduction to path analysis. *Canadian Journal of Psychiatry*, *50*(2), 115–122.
Tekkumru-Kisa, M., Stein, M. K., & Schunn, C. (2015). A framework for analyzing cognitive demand and content-practices integration: Task analysis guide in science. *Journal of Research in Science Teaching*, *52*(5), 659–685.
Tytler, R. (2007). *Re-imagining science education: Engaging students in science for Australia's future*. Retrieved from Camberville, VIC.
Wu, H.-K., & Hsieh, C. E. (2006). Developing sixth graders' inquiry skills to construct explanations in inquiry-based learning environments. *International Journal of Science Education*, *28*(11), 1289–1313.
Wu, P.-H., Wu, H.-K., & Hsu, Y.-S. (2014). Establishing the criterion-related, construct, and content validities of a simulation-based assessment of inquiry abilities. *International Journal of Science Education*, *36*(10), 1630–1650.
Zhang, W.-X., Hsu, Y.-S., Wang, C.-Y., & Ho, Y.-T. (2015). Exploring the impacts of cognitive and metacognitive prompting on students scientific inquiry practices within an e-learning environment. *International Journal of Science Education*, *37*(3), 529–553.

Appendix 1

Conceptual knowledge scores in pre and post unit tests.

			Conceptual knowledge					
						Wilcoxon signed-rank test		
	(*N* = 117)		Total score	*M*	*SD*	*Z*	*p*	*ES*[a]
Structured inquiry	Evolution	Pre	32	8.81	4.21	7.13	<.001	0.66
		Post		12.14	4.87			
	Buoyance	Pre	11	5.66	2.48	5.22	<.001	0.48
		Post		6.84	2.00			
	Plate tectonics	Pre	8	3.42	1.32	8.36	<.001	0.77
		Post		5.73	1.53			
Guided inquiry	Genetics	Pre	20	9.03	3.10	7.48	<.001	0.69
		Post		12.18	3.03			
	Temperature and heat	Pre	9	6.06	1.43	2.40	.016	0.22
		Post		6.40	1.54			
	Dam	Pre	17	9.50	3.31	7.13	<.001	0.66
		Post		11.50	3.11			

[a]Effect size = $Z/\sqrt{N}$.

Total IA scores in pre and post unit tests.

			Inquiry ability					
						Wilcoxon signed-rank test		
	(*N* = 117)		Total score	*M*	*SD*	*Z*	*p*	*ES*[a]
Structured inquiry	Evolution	Pre	30	12.43	4.47	5.25	<.001	0.49
		Post		14.30	4.41			
	Buoyance	Pre	27	8.92	4.99	8.92	<.001	0.82
		Post		14.88	5.10			
	Plate tectonics	Pre	25	6.74	2.92	9.01	<.001	0.83
		Post		15.87	5.57			
Guided inquiry	Genetics	Pre	17	10.45	2.56	7.21	<.001	0.67
		Post		12.57	2.42			
	Temperature and heat	Pre	19	10.82	3.98	3.47	.001	0.32
		Post		11.94	3.79			
	Dam	Pre	25	8.14	6.59	7.60	<.001	0.70
		Post		14.58	5.75			

[a]Effect size = $Z/\sqrt{N}$

The students' total scores during the units (immediate assessments).

		(*N* = 117)	
	Total score[a]	*M*	*SD*
Evolution	50	25.50	8.94
Buoyance	38	23.57	5.38
Plate tectonics	46	29.52	6.72
Genetics	88	59.54	11.69
Temperature and heat	63	41.39	8.27
Dam	51	30.79	6.95

[a]Total score = conceptual item score + IA item score.

Appendix 2

Direct effects in factors influencing 'conceptual knowledge demonstrated in the proximal conceptual knowledge assessment' (integrated knowledge built).

Variables	Endogenous (dependent variables)		
	Learning outcomes conceptual knowledge	Learning outcomes IA	Integrated knowledge built
Prior knowledge	.50*	.07	.42*
Prior IA	.25*	.05	−.04
Learning outcomes conceptual knowledge	–	.82*	.02
Learning outcomes IA	–	–	.50*
R^2	.46	.79	.67

*$p < .05$.

Direct, indirect and total *significant* effects on 'conceptual knowledge demonstrated in the proximal conceptual knowledge assessment' (integrated knowledge built).

Variables	Direct	Indirect	Total
Prior knowledge	.42	.20	.62
Prior IA		.10	.10
Learning outcomes conceptual knowledge		.41	.41
Learning outcomes IA	.50	–	.50

Appendix 3

Direct effects in factors influencing 'conceptual knowledge and IA demonstrated in *SIU* post-tests'.

Variables	Endogenous (dependent variables)			
	Knowledge demonstrated during class	IA demonstrated during class	Knowledge demonstrated in unit post-test	IA demonstrated in unit post-test
Prior knowledge (unit pre-test)	.20*	.25	.32*	.09
Prior IA (unit pretest)	−.12	.28*	.07	.27*
Knowledge demonstrated during class	–	–	.37*	.47*
IA demonstrated during class	.81*	–	.18	.10
R^2	.74	.35	.64	.65

*$p < .05$.

Direct, indirect, and total *significant* effects on '*conceptual knowledge* demonstrated in *SIU post-tests*'.

Variables	Direct	Indirect	Total effects
Prior knowledge (unit pretest)	.32	.07	.39
Prior IA (unit pretest)		.11	.11
Knowledge demonstrated during class	.37		.37
IA demonstrated during class		.29	.29

Direct, indirect, and total *significant* effects on '*IA* demonstrated in *SIU post-tests*'.

Variables	Direct	Indirect	Total effects
Prior knowledge (unit pretest)		.10	.10
Prior IA (unit pretest)	.27	.14	.41
Knowledge demonstrated during class	.47	–	.47
IA demonstrated during class		.38	.38

Appendix 4

Direct effects in factors influencing ‘conceptual knowledge and IA demonstrated in *GIU* post-tests’.

	Endogenous (dependent variables)			
Variables	Knowledge demonstrated during class	IA demonstrated during class	Knowledge demonstrated in unit post-test	IA demonstrated in unit post-test
Prior knowledge (unit pretest)	.18	−.01	.53*	.17*
Prior IA (unit pretest)	.49*	.02	.02	.15
Knowledge demonstrated during class	–	.86*	.03	.19
IA demonstrated during class	–	–	.28*	.36*
R^2	.38	.76	.54	.55

*$p < .05$.

Direct, indirect and total *significant* effects on ‘*conceptual knowledge* demonstrated in *GIU post-tests*’.

Variables	Direct	Indirect	Total effects
Prior knowledge (unit pretest)	.53		.53
Prior IA (unit pretest)		.12	.12
Knowledge demonstrated during class		.24	.24
IA demonstrated during class	.28		.28

Direct, indirect, and total *significant* effects on ‘*IA* demonstrated in *GIU post-tests*’.

Variables	Direct	Indirect	Total effects
Prior knowledge (unit pretest)	.17		.17
Prior IA (unit pretest)		.15	.15
Knowledge demonstrated during class		.31	.31
Inquiry ability demonstrated during class	.36		.36