Instructional Science (2020) 48:79–113
https://doi.org/10.1007/s11251-020-09503-8

ORIGINAL RESEARCH

Constructing a model of engagement in scientific inquiry: investigating relationships between inquiry-related curiosity, dimensions of engagement, and inquiry abilities

Pai-Hsing Wu[1,2] · **Hsin-Kai Wu**[2,3]

Received: 21 May 2019 / Accepted: 20 January 2020 / Published online: 4 February 2020

Abstract
According to policy documents and research studies, one key objective of science education is to develop students' inquiry abilities; however, relatively little is known about the interplay among students' inquiry abilities, the dimensions of their engagement, and their inquiry-related curiosity. The purpose of this study is to explore how four dimensions of engagement (i.e., cognitive, behavioral, emotional, and social) were driven by inquiry-related curiosity and how they affected the students' inquiry abilities. Structural equation modeling was employed to analyze data collected from 605 11th graders, including their responses to items in an online questionnaire and their performances on a computer-based assessment of scientific inquiry abilities. The results showed that students' curiosity was associated with their inquiry abilities, and such an association was partially mediated by the four dimensions of engagement in science laboratory classes. Moreover, the results revealed that among the four dimensions of engagement, only cognitive and emotional engagement had significant total effects on students' inquiry abilities and that the influence of behavioral and social engagement on inquiry abilities was completely mediated by cognitive engagement. This study suggests a critical role played by emotional engagement, cognitive engagement, and curiosity in developing students' inquiry abilities.

Keywords Behavioral engagement · Cognitive engagement · Curiosity · Emotional engagement · Inquiry ability · Social engagement

✉ Hsin-Kai Wu
hkwu@ntnu.edu.tw

[1] Department of Technology Application and Human Resource Development, National Taiwan Normal University, Taipei, Taiwan

[2] Graduate Institute of Science Education, National Taiwan Normal University, P.O. Box 97-27, Taipei 11699, Taiwan

[3] Faculty of Education, University of Johannesburg, Johannesburg, South Africa

Introduction

For the last two decades, engagement has been conceived as an important factor affecting students' learning and academic success (Fredricks et al. 2004; Kuh 2003; Newmann et al. 1992) and refers to the observable and unobservable qualities of students' involvement and participation in learning activities (Ryan and Patrick 2001). Many researchers have suggested that the qualities of students' participation in learning activities could involve observable behaviors (Zumbrunn et al. 2014), internal cognitions (Greene et al. 2004), as well as personal emotions (Fredricks et al. 2004) in school contexts. These studies have supported a multifaceted conceptualization of engagement, and four dimensions of engagement have been identified, namely behavioral engagement, emotional engagement, cognitive engagement, and social engagement (Fredricks et al. 2016a, b). Although previous research has provided evidence for the effects of every single dimension or a set of dimensions on students' learning, the understanding of the relationships among these dimensions of engagement with respect to learning outcomes is very limited. Additionally, engagement in subject areas has recently been receiving increasing attention (e.g., Wang et al. 2016). In science education, scientific inquiry has been considered as a valuable and inseparable part of science teaching and learning (National Research Council [NRC] 1996, 2000) but relatively little is known about the role of students' engagement in their inquiry performances. Therefore, to construct a model of engagement in scientific inquiry, this study examined the interplay among the four dimensions of students' engagement, inquiry-related curiosity, and inquiry abilities of students.

Why should inquiry-related curiosity be considered as a potential predictor of engagement? Engagement could be affected by motivation-based factors, such as belonging, self-efficacy (Zumbrunn et al. 2014), and intrinsic motivation (Froiland and Worrell 2016), and curiosity (Wu et al. 2018). Among these factors, curiosity has been valued and encouraged in science education because students' scientific curiosity about how the world works or why a phenomenon happens may motivate them to conduct investigations and to find these answers through scientific inquiry (NRC 2000, 2012). Curiosity can be defined as a psychological disposition for the unknown, novelty, ambiguity, and uncertainty (Litman 2005; Mussel 2010) and motivate students to pursue novel information and challenging experiences (Sinha et al. 2017). Additionally, Luce and Hsi (2015) argued for "science-relevant curiosity" such as wonderment about casual mechanisms and inconsistent observations in science. Students' curiosity about science phenomena and inquiry activities could "develop into a deeper, sustained, persistent pursuit of scientific reasoning and science learning" (p. 74) and then lead to higher engagement. For example, when students have wonderment about a phenomenon, a knowledge gap between what they already know and what they want to know may occur (Chin and Osborne 2008). Such a knowledge gap could motivate students to explore how and why the phenomenon happens and to demonstrate inquiry behaviors. Thus, inquiry-related curiosity was examined in this study because such a psychological state and disposition could facilitate students' engagement in authentic and meaningful science learning.

On the other hand, developing students' inquiry abilities has been emphasized as a key objective of science learning in many policy documents (NRC 2000) and research studies (Krajcik et al. 1998). These fundamental abilities enable students to coordinate science process skills and knowledge, to conduct meaningful scientific investigations, and to successfully participate in scientific inquiry. This study thus focuses on students' abilities in four inquiry activities including asking scientific questions, planning experiments, analyzing data, and formulating scientific explanations. However, numerous empirical studies have indicated that students usually lack these fundamental abilities and encounter great difficulties in the process of inquiry (e.g., Bell 2002; Jeong et al. 2007). To understand whether and how qualities of students' participation in inquiry activities could enhance their abilities, this study aims at exploring how the four dimensions of engagement (i.e., cognitive, behavioral, emotional, and social) are driven by inquiry-related curiosity and how they affect inquiry abilities. To accomplish this purpose, we employed structural equation modeling to analyze data collected from 605 11th graders, including their responses to items in an online questionnaire and their performances on a computer-based assessment of scientific inquiry abilities.

Theoretical and empirical underpinnings

Definition of engagement

Engagement has been viewed as an important construct in educational research because it could influence students' learning processes and outcomes (Astin 1984; Engle and Conant 2002; Finn 1989). Although it is commonly accepted that engagement can be regarded as the qualities of students' involvement and participation in learning activities (Ryan and Patrick 2001), different aspects of engagement have been introduced for various tasks, purposes, and contexts. For example, Astin (1984) advocated the notion of student involvement to investigate students' energy and time investment in both physiological and psychological aspects. On the other hand, Engle and colleagues adopted the notion of disciplinary engagement to illustrate students' involvement in the particular subject matter in classroom settings (Engle and Conant 2002; Forman et al. 2014). Furthermore, engagement can be used to describe and explain students' behaviors and participation in schools. With a higher degree of school engagement, students would demonstrate more willingness to search for the strategies to perform better (Finn 1989) and show less anxiety and resistance to school activities (Archambault et al. 2009).

To highlight and address the different aspects of engagement, researchers have argued for a multi-dimensional definition of engagement (Archambault et al. 2009; Newmann et al. 1992). Four dimensions identified by previous research include cognitive engagement, behavioral engagement, emotional engagement, and social engagement. Cognitive engagement refers to students' proficiency of tasks, use of strategies, or pursuit of achievement to satisfy the requirements of learning tasks (Greene et al. 2004). Behavioral engagement involves students' attentiveness, persistence, and investment of time (Kuh 2009; Newmann et al. 1992). Emotional engagement refers to students' perceptions, values or feelings about learning activities and environments (Fredricks et al. 2004). Social engagement refers to students' interactions with their peers, teachers or staff members in schools (Järvelä et al. 2016). These four dimensions allow a comprehensive and analytical look at students' participation in learning.

Relationships among the dimensions of engagement

As different dimensions of the same construct, behavioral, cognitive, emotional, and social engagement could interact with each other, investigations of the interplay between the dimensions could assist researchers and educators in understanding how to facilitate a certain dimension of engagement. Archambault et al. (2009) developed a questionnaire to assess the behavioral, cognitive, and emotional engagement of students, and conducted model comparisons to investigate the relationships among them. Although they found that "the three dimensions of engagement were highly covariant" (p. 665), especially the emotional and cognitive dimensions, the study did not illustrate how the dimensions of engagement affect each other. On the other hand, a longitudinal study done by Li and Lerner (2013) indicated that "behavioral and emotional engagement were related bidirectionally" (p. 20) and that cognitive engagement could be predicted by behavioral engagement. In addition, by comparing results from Grades 9, 10, and 11, Lin and Lerner found a significant pathway from the emotional engagement at Grade 10 to cognitive engagement at Grade 11. Gunuc and Kuzu (2015) reconfirmed the predictive link of behavioral engagement to cognitive engagement, and further revealed that emotional engagement could affect behavioral engagement in a technology-based learning environment. Moreover, in a collaborative learning environment, students' emotions could influence their social and behavioral engagement in small groups (Linnenbrink-Garcia et al. 2011); for example, positive emotions were positively associated with positive group interactions (e.g., actively working to support group members' engagement).

Taken together, the aforementioned studies suggest a hypothesis whereby emotional engagement could predict behavioral and social engagement, and may directly or indirectly influence cognitive engagement. These relationships constitute a central part of our hypothesized model (Fig. 1) and so far relatively little research has directed its attention to these relationships in disciplinary engagement. Furthermore, this study addresses the issue of how the dimensions of engagement affect students' performances in scientific inquiry.

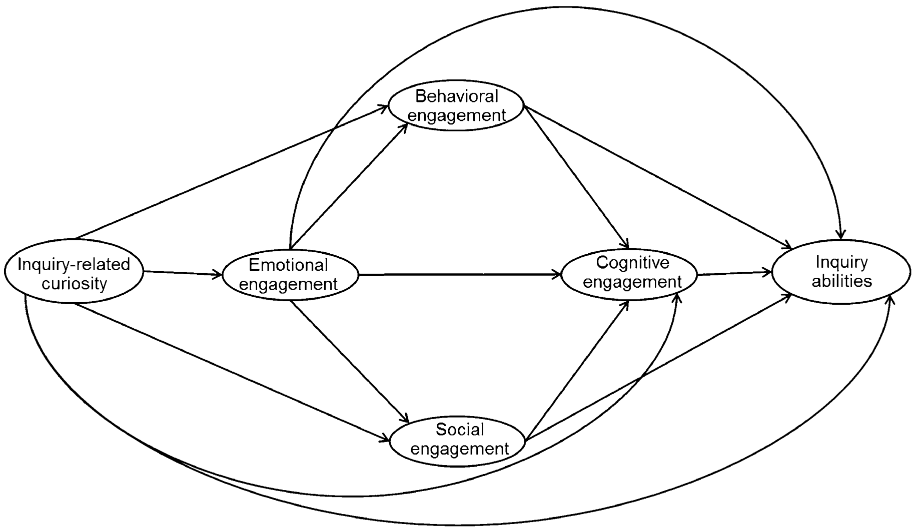

Fig. 1 A hypothesized model of relationships among students' inquiry-related curiosity, dimensions of engagement, and inquiry abilities

Engagement and student performances

Previous research has investigated how engagement could affect students' learning, cognitive, and school performances. Among the four dimensions, cognitive engagement has been the most investigated and its positive association with learning achievement has been well documented (e.g., Greene et al. 2004). The explanation of the association is that cognitive engagement involves the use of learning strategies and cognitive processing, which could result in different levels of achievement, such as test scores (Greene et al. 2004) and literature achievement (Nystrand and Gamoran 1991). Other dimensions of engagement could also predict students' school performances. Archambault et al. (2009) found that the overall engagement had a significant but negative predictive link to school dropout and that their further examination showed that only students' behavioral engagement was significantly associated with school dropout. Additionally, in their mixed-methods research, Wu and Huang (2007) revealed that students' emotional engagement in computer-based science learning had no impact on their achievement. Yet, the findings from Authors' qualitative data analyses suggested interactions among students' achievement levels, cognitive engagement, and behavioral engagement. For instance, the high-achieving group was more engaged in the cognitive and behavioral dimensions.

Although the association between engagement and learning performances has been confirmed by previous research, it is still unclear whether social engagement affects students' learning performances and how the different dimensions of engagement mediate or influence the relationships between engagement and learning performances. Therefore, situated in the context of disciplinary engagement, this study takes the four dimensions of engagement into consideration and hypothesizes that they could positively affect students' inquiry abilities directly and indirectly (Fig. 1).

Definition and framework of inquiry abilities

Inquiry has been viewed as an essential component of science learning. Scientific inquiry refers to multifaceted activities similar to what scientists do to construct understandings about the world, including questioning, experimenting, designing, planning, analyzing, interpreting, explaining, arguing, and communicating (Barrow 2006; Duschl et al. 2007; Krajcik et al. 1998; NRC 1996, 2000; Pedaste et al. 2015). Fundamental abilities that combine knowledge with skills are required to undertake these inquiry activities (NRC 2000). These abilities are defined as inquiry abilities in this study, and have been emphasized as key learning objectives of science education in policy documents internationally (e.g., NRC 2000; Ministry of Education [MOE] 2018).

To develop an assessment framework that involves the nature of inquiry and reflects the process of inquiry, we reviewed science education standards (NRC 1996, 2000, 2012), science curriculum guidelines (MOE 2018), and research studies (e.g., Wu and Huang 2007; Wu et al. 2015; Krajcik et al. 1998; Pedaste et al. 2015). These documents presented a number of frameworks for science inquiry. One of the most recent and influential frameworks was proposed by the National Generation Science Standards (NRC 2012), in which three spheres of activity for scientists were identified: Investigating (asking questions, making observations and doing experiments), Evaluating (arguing, critiquing, and analyzing), and Developing Explanations (developing theories and models and formulating hypotheses). Additionally, by reviewing 32 articles on inquiry phases and cycles, Pedaste et al.

(2015) provided a synthesized framework that is constituted of five distinct general inquiry phases: Orientation, Conceptualization (questioning and hypothesis generation), Investigation (exploration, experimentation, and data interpretation and analysis), Conclusion, and Discussion (communication and reflection). The two frameworks reflect the multifaceted nature of inquiry and cover a variety of inquiry activities. Although they categorized some inquiry activities into different phases or spheres, four inquiry activities that were identified in both frameworks are asking questions, conducting investigations, analyzing data, and developing explanations and conclusions. These four major inquiry activities thus were selected as the backbone of our framework.

Our assessment framework of inquiry abilities focuses on four main abilities and related sub-abilities: (1) Questioning (sub-abilities: formulating questions, identifying questions, and making predictions); (2) Experimenting (sub-abilities: identifying controlled and manipulated variables, planning experimental procedures, and selecting appropriate measurements); (3) Analyzing (sub-abilities: identifying relevant data, and transforming data); and (4) Explaining (sub-abilities: making a claim, using evidence, reasoning from evidence to the claim, and offering and evaluating alternative explanations). We also developed and validated a computer-based assessment that integrated animations, videos, pictures, and simulations into the test and task design to measure students' inquiry abilities (Wu et al. 2015; Kuo et al. 2015).

Inquiry abilities, engagement, and inquiry-related curiosity

Although inquiry abilities are foundational to science learning, previous research has indicated that students encounter difficulties developing the abilities. Regarding the questioning ability, for example, students may not be able to express ideas and questions clearly (Krajcik et al. 1998), and the questions raised by students could be too divergent to predict, and cannot be answered by experimental results (Apedoe 2008). When doing experiments, students may collect ineffective experimental data because they lack the ability to control variables (Krajcik et al. 1998). Additionally, some middle school students do not have the ability to appropriately organize and interpret data (Jeong et al. 2007). They may not understand the components of scientific explanations (e.g., claim, reasoning, and evidence) and may not be able to use data as evidence to support their claims (McNeill and Krajcik 2008).

To explore possible ways to foster students' inquiry abilities, researchers have identified factors that may associate with the development of students' inquiry abilities (Wu et al. 2018; Luce and Hsi 2015). By examining data collected from more than 2000 eighth and 11th graders, Wu et al. (2018) revealed a positive relationship between students' engagement in school laboratory activities and their inquiry abilities. However, Wu et al. (2018) focused only on students' behavioral engagement in laboratory activities and did not provide information about how other dimensions of engagement could influence inquiry abilities.

Furthermore, Wu et al. (2018) indicated that students' scientific curiosity could predict their inquiry abilities. Curiosity can be defined as a psychological disposition for the unknown, novelty, ambiguity, and uncertainty (Litman 2005; Mussel 2010) and measured as "the threshold of desired uncertainty in an environment which leads to exploratory behavior" (Jirout and Klahr 2012, p. 150). Luce and Hsi (2015) argued for the importance of characterizing students' curiosity expressions with respect to scientific inquiry, and the relationship between curiosity and engagement in inquiry activities has been suggested in previous studies (e.g., Callanan and Jipson 2001; Callanan and Oakes 1992; Engel 2011). One possible explanation is that students' science curiosity could initiate them into questioning and exploration (Chin and Osborne 2008) and drive their engagement in inquiry activities such as explaining and experimenting; this, in turn, may facilitate the development of students' inquiry abilities. Thus, our hypothesized model in Fig. 1 also includes the potential impact of inquiry-related curiosity on the dimensions of engagement and inquiry abilities.

Research questions and proposed model

The purpose of this study is to examine how the four dimensions of engagement (i.e., cognitive, behavioral, emotion, and asocial engagement) are driven by students' inquiry-related curiosity, and how they affect their inquiry abilities. Four research questions guided this study.

(1) What are the relationships among the four dimensions of engagement?
(2) What are the effects of the dimensions of engagement on inquiry abilities?
(3) Is students' inquiry-related curiosity associated with the dimensions of engagement?
(4) Do the dimensions of engagement mediate the association between students' inquiry-related curiosity and their inquiry abilities?

On the basis of the theoretical and empirical underpinnings, we proposed the model in Fig. 1 and formulated hypotheses for the research questions as follows. (1) Emotional engagement predicts behavioral and social engagement, and directly or indirectly affects cognitive engagement (Archambault et al. 2009; Li and Lerner 2013; Linnenbrink-Garcia et al. 2011). (2) The four dimensions of engagement positively influence inquiry abilities in a direct or indirect manner (Wu and Hang 2007; Greene et al. 2004). (3) Inquiry-related curiosity has a positive impact on the four dimensions of engagement (Wu et al. 2018; Engel 2011). (4) The dimensions of engagement play a mediating role in the association between students' inquiry-related curiosity and their inquiry abilities (Chin and Osborne 2008). By examining the relationships among high school students' curiosity, engagement and inquiry abilities, this study could provide insight into how to foster students' inquiry abilities by enhancing different dimensions of engagement and promoting students' curiosity.

Methods

Sampling

To achieve the purpose of this study, we recruited 11th grade students from senior high schools in three northern cities of Taiwan to participate in this study. The sample size was determined based on the suggestion of Westland (2010). The appropriate sample size in structural equation modeling should consider both the ratio of indicators to latent variables and the minimum effect in SEM at a given significance and power. For the structural model with six latent variables and 40 indicators in this study (1 ability, 32 engagement, 5 curiosity, and 2 computer experience indicators), according to Westland (2010), the suggested sample size was 675 for power = 0.80, $\alpha = 0.05$, and the medium effect size (R^2) = 0.15.

Using a two-stage stratified cluster design, we first divided the 78 senior high schools of the three cities into eight strata based on their students' percentile ranks in the Basic Competence Test for Junior High School Students. In the second stage, we sampled one school from each stratum. The class size at high schools in Northern Taiwan was usually between 40 and 45 so we invited two classes in each sampled school (a total of 16 classes) to achieve the suggested sample size of 675. Yet, some classes were smaller than we expected, and the 16 classes with the total of 647 students were recruited to participate in this study. Among the students, 42 offered no responses to the administered questionnaires; their data were therefore excluded, resulting in a sample of 605 11th graders.

Measures

Three instruments were used to collect data in this study: a computer-based assessment of scientific inquiry abilities (39 items), a 32-item questionnaire of engagement, and a questionnaire for collecting information of students' inquiry-related curiosity, computer experience, and socioeconomic status. Table 1 shows the descriptions of indicators for the latent variables in this study.

Inquiry abilities

A computer-based assessment was developed to evaluate high school students' questioning, experimenting, analyzing, and explaining abilities (Kuo et al. 2015). The assessment used scenario-based tasks to address the four inquiry abilities and incorporated simulations, animations, and videos to engage students in meaningful inquiry situations. For example, in the Camera Task (see Fig. 2), a simulated camera was provided and allowed students to change the aperture range and shutter speed, to observe the brightness of the photo, and to investigate the relationships between variables. The item pool of the assessment contained 101 items in 26 tasks which were developed to examine students' inquiry abilities across four science content areas (i.e., chemistry, physics, biology, and earth science).

The validity and reliability of the assessment items were established by our previous studies (Wu et al. 2014; Kuo et al. 2015). In this study, we selected 39 items from the pool to assess students' four inquiry abilities: 11 items for questioning, 10 for experimenting, 10 for analyzing, and 8 for explaining. Three professors of science education were invited to evaluate the items for content validity. The content validity among the three experts was excellent (kappa = 0.88–0.96) and indicated their agreement that these tasks were sufficiently relevant to inquiry abilities.

Moreover, scoring rubrics were developed to evaluate the open-ended items. These rubrics included the level codes to indicate the performance levels in our framework, the score codes for actual points, and the examples of anticipated responses (Table 2). Three experienced researchers who had joined our previous studies were invited to score the students' responses. The scoring researchers were trained to establish interrater reliability. The process involved two runs of grading 60 students' responses to all 39 items, and the inter-rater agreements in scoring students' open-ended answers ranged from 81.2 to 95.0%. After the agreements were achieved, each rater rated the students' responses by using the scoring rubrics. Student responses from the 39 items were scored and gathered. Unidimensional Rasch models were used for analysis to derive five sets of plausible values (PVs) as the indicator of students' inquiry abilities (Wu and Adams 2007). By doing so, the students' inquiry abilities were accurately estimated (Wu et al. 2016). The reliability of expected a

Table 1 Descriptions of indicators for the latent variables

Indicators	Description
Cognitive engagement	
CE_1	I go through the work for laboratory classes and make sure that it's right
CE_2	I think about different ways to solve a problem
CE_3	I try to connect what I am learning to things I have learned before
CE_4	I try to understand my mistakes when I get something wrong
CE_5	I would rather be told the answer than have to do the work
CE_6	I don't think that hard when I am doing work for class
CE_7	When work is hard, I only study the easy parts
CE_8	I try to plan an approach in my mind before I actually start homework or studying
CE_9	I try to put the ideas in my own words when learning new information
Behavioral engagement	
BE_1	I stay focused on enacting experiments
BE_2	I put effort into my experiments
BE_3	I keep trying even if something is hard
BE_4	I complete my homework on time
BE_5	I talk about science/math outside of class
BE_6	I don't participate in class
BE_7	I do other things when I am supposed to be paying attention
BE_8	If I don't understand, I give up right away
Emotional engagement	
EE_1	I look forward to laboratory classes
EE_2	I enjoy learning new things during laboratory classes
EE_3	I want to understand what is learned in laboratory classes
EE_4	I feel good when I am in laboratory classes
EE_5	I think that laboratory classes are boring
EE_6	I don't want to be in laboratory classes
EE_7	I don't care about learning in laboratory classes
EE_8	I often feel down when I am in laboratory classes
EE_9	I get worried when I learn new things in laboratory classes
Social engagement	
SE_1	I build on others' ideas
SE_2	I try to understand other people's ideas in laboratory classes
SE_3	I try to work with others who can help me in laboratory classes
SE_4	I try to help others who are struggling in laboratory classes
SE_5	I don't care about other people's ideas about the experiment
SE_6	When working with others, I don't share ideas
Inquiry-related curiosity	
Curi1	The degree to which students are curious about whether their own proposed research questions are feasible
Curi2	The degree to which students are curious about the rationale of experiment design
Curi3	The degree to which students are curious about how to interpret experiment results
Curi4	The degree to which students are curious about the possible interpretations of conflicts between experiment results and hypotheses

Table 1 (continued)

Indicators	Description
Curi5	The degree to which students are curious about the possible explanations of experiment data

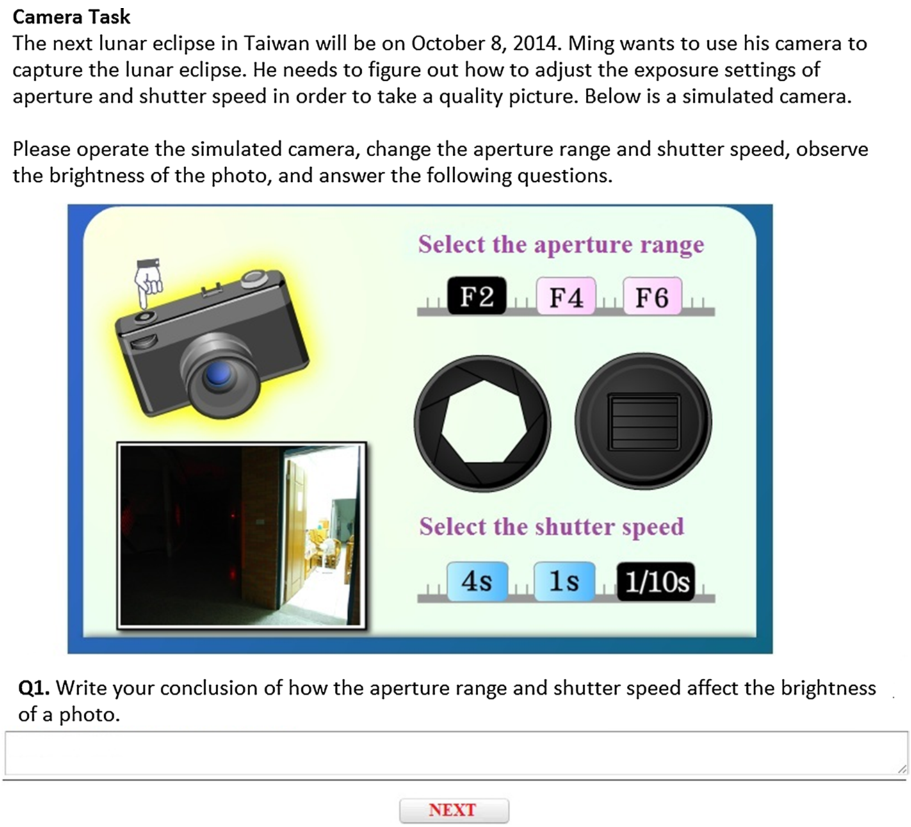

Fig. 2 The camera task in the computer-based assessment of scientific inquiry abilities

posteriori estimation based on the plausible values (EAP/PV) was 0.88, which showed that the test could explain a high percentage of the variation in the students' inquiry abilities.

Engagement in scientific inquiry

This study adapted items designed by Fredricks and her colleagues (Fredricks et al. 2016a, b; Wang et al. 2016) to measure the four dimensions of students' engagement in scientific inquiry activities. In the questionnaire, the learning context was described as "science laboratory classes" because the students may not be familiar with the term "scientific inquiry" and the learning opportunities of scientific inquiry have presumably been offered in school laboratory classes (Hofstein and Lunetta 2004).

Table 2 The scoring rubrics of making a claim for the camera task

Test item	After operating the simulated camera by changing the aperture range and shutter speed, please write down how the aperture range and shutter speed affect the brightness of a photo		
Level code	Performance	Score point	Description and example
High	Students generate correct claims and the claims completely articulate the relationships between the variables and the trends of the data	2	Description: A response includes correct claims about how the aperture range and shutter speed affect the brightness of a photo Example: the narrower the aperture range is and the faster the shutter speed is, the darker the photo is (*including correct claims about both relationships*)
Middle	Students generate correct claims but the claims incompletely articulate the relationships between the variables and the trends of the data	1	Description: A response includes correct claims about how the aperture range or shutter speed affects the brightness of a photo Example: When the shutter speed is kept at 4 s, the wider the aperture range is, the brighter the photo is (*only including a correct claim about one relationship*)
Low	Students identify or generate incorrect claims	0	Description: A response includes incorrect claims about how the aperture range or the shutter speed affects the brightness of a photo or consists only of irrelevant information Example: The wider the aperture range is and the faster the shutter speed is, the more difficulty I have taking a photo in the dark (*including a claim that is irrelevant to the brightness of a picture*)

The participating students were asked to reflect on the degree of their behavioral, cognitive, emotional and social interactive states in science laboratory classes (see the questions in Table 1). Nine indicators were used to evaluate their cognitive engagement (CE) such as thinking about different ways to solve a problem. Eight indicators were to gather information about behavioral engagement (BE), including items about attention, on-task behaviors, and efforts. In the part of emotional engagement (EE), nine indicators included positive and negative emotions, interest, and perceptions of the value of the science laboratory. Six indicators evaluated their social engagement (SE) with peers such as understanding different perspectives, building on others' ideas, and helping peers. Students rated each item from *never* (1) to *always* (4).

A pilot study was conducted to validate the questionnaire and a high school located in northern Taiwan (but outside of the three northern cities where the formal study was conducted) participated in the study. All of the 11th graders in the school were invited to answer the questionnaire, and data from 114 students were collected. The Cronbach's alpha for the four dimensions of engagement ranged from 0.80 to 0.85, and the Cronbach's alpha of all 32 items was 0.94. The reliability results suggested that these items provided a high quality of internal consistency to evaluate students' engagement in laboratory classes. Moreover, the confirmatory factor analysis (CFA) showed the composite reliabilities (CR) of all items ranging between 0.83 to 0.85 (higher than 0.60), which demonstrated a sufficient internal consistency of the latent variables (Bagozzi and Yi 1988). The final version contained 32 items and had suitable reliability and validity to measure students' inquiry-related engagement.

Inquiry-related curiosity

We developed five items to measure students' inquiry-related curiosity (Table 1). Students were asked to rate the five statements on a 4-point scale, from *never* (1) to *always* (4). These items probed the degree to which the students were curious about the feasibility of their research questions, the rationale of the experiment design, and the possible interpretations and explanations of the experiment data. Our previous study validated these five items (Wu et al. 2018) and suitable reliability was obtained ($\alpha = 0.923$). In this study, reliability was reexamined and the value of composite reliability was 0.93 (Table 3), indicating a high internal consistency for the measure.

Computer experience

Because the questionnaires and assessment were administered on computers, students' computer experience and familiarity with the computer interface could influence their performances (Scheuermann and Björnsson 2009). To focus on the relationships in our research questions, in this study, students' computer experience was measured and later controlled in the data analysis. Two items were developed to assess the students' computer experience. In the first item, students reported the frequency with which they used computer devices outside of school from *never* (1) to *everyday* (5). In the second item, students indicated whether web browsing, among three other computer activities (i.e., social networking, online gaming, and word processing), was the most frequent computer activity that they engaged in. Dummy coding was used for this item (engaged = 1; not engaged = 0).

Table 3 Parameters of the initial and re-specified measurement models

Latent variable	Indicator	Initial model	Final specified model			
		Standardized factor loading (SE)	Standardized factor loading (SE)	CR	AVE	Residual variance
Inquiry-related curiosity (Curi)	Curi1	0.74 (0.060)***	0.74 (0.060)***	.93	0.72	0.45
	Curi2	0.77 (0.060)***	0.77 (0.060)***			0.41
	Curi3	0.81 (0.060)***	0.81 (0.060)***			0.34
	Curi4	0.83 (0.060)***	0.83 (0.060)***			0.31
	Curi5	0.76 (0.060)***	0.76 (0.060)***			0.42
Cognitive engagement (CE)	CE_1	0.60 (0.090)***	0.60 (0.090)***	.83	0.46	0.64
	CE_2	0.72 (0.090)***	0.72 (0.090)***			0.48
	CE_3	0.72 (0.090)***	0.72 (0.090)***			0.48
	CE_4	0.73 (0.090)***	0.73 (0.090)***			0.47
	CE_8	0.61 (0.090)***	0.61 (0.090)***			0.63
	CE_9	0.67 (0.100)***	0.67 (0.100)***			0.55
Behavioral engagement (BE)	BE_1	0.91 (0.030)***	0.91 (0.030)***	.76	0.91	0.17
	BE_2	0.94 (0.030)***	0.95 (0.030)***			0.10
	BE_3	0.76 (0.040)***	0.75 (0.040)***			0.44
	BE_4	0.54 (0.050)***	–[a]			–[a]
	BE_7	0.35 (0.040)***	–[a]			–[a]
	BE_8	0.29 (0.040)***	–[a]			–[a]
Emotional engagement (EE)	EE_1	0.83 (0.040)***	0.82 (0.040)***	.91	0.71	0.33
	EE_2	0.93 (0.040)***	0.94 (0.040)***			0.12
	EE_3	0.88 (0.030)***	0.89 (0.040)***			0.21
	EE_4	0.71 (0.040)***	0.71 (0.040)***			0.50
	EE_5	0.44 (0.040)***	–[a]			–[a]
	EE_6	0.45 (0.040)***	–[a]			–[a]
	EE_7	0.46 (0.030)***	–[a]			–[a]
Social engagement (SE)	SE_1	0.77 (0.060)***	0.77 (0.060)***	.80	0.51	0.41
	SE_2	0.77 (0.060)***	0.77 (0.060)***			0.41
	SE_3	0.60 (0.050)***	0.60 (0.050)***			0.64
	SE_4	0.69 (0.050)***	0.69 (0.050)***			0.52

CR composite reliability, *AVE* average variance extracted

[a]The value was not estimated because the indicator was deleted due to a loading lower than 0.55

***$p < .001$

Socioeconomic status

To determine the relationships between curiosity, engagement, and inquiry abilities, this study also measured and controlled students' socioeconomic status because it has been identified as an influential factor affecting students' academic performances (May 2006). Four items were used to assess socioeconomic status. Two items asked students to indicate their parents' education levels from no education (0) to having graduate degrees (6). One item was to estimate their family income according to seven income ranges (coded as 1–7).

The last item reported the availability of home possessions and family conditions including rooms owned by students, cars, video recorders, two or more televisions, two or more digital computing devices (i.e., desktop, laptop and tablet computers), and mobile phones owned by their families. The z-scores of the four items were calculated and summed up as a SES variable.

Data analysis

Structural equation modeling

We took a two-step modeling approach (Anderson and Gerbing 1988) to test our hypothesized model. In the first step, CFA was employed to examine construct validity and to investigate whether the hypothesized relationships between indicators and the associated latent variables in our measurement model were valid. The relationships among items of the same latent variable were examined by convergent validity, and the factor loading must be larger than 0.55 (Comrey and Lee 1992; Harrington 2008). Furthermore, the reliability was demonstrated by composite reliability coefficients with a satisfactory value greater than 0.60 (Fornell 1982). The minimum cut-off of average variance extracted (AVE) for each latent variable was 0.50 to indicate that a greater common variance is captured by a construct than the amount of variance due to measurement error (Bagozzi and Yi 1988; Fornell 1982).

In the second step, the latent relationships in our structural model were tested using structural equation modeling with the validated latent variables. To avoid the confounding effect of students' computer experiences and socioeconomic status, the two variables were used as controlling variables in the models to control all of the examined effects above and beyond the effects of the two variables.

This study also used measures suggested by previous research to evaluate the goodness of fit of our structural model. For the $\frac{\chi^2}{df}$ ratio, as indicated by Bollen (1989), "there is no consensus on what represents a 'good' fit, with recommendations ranging from ratios of 3, 2 or less […] to as high as 5" (p. 278). We took the recommendation from Kelloway (1998) and West et al. (2012) set 5 as the cut-off value. Other indications of good fit used in this study included a root mean square error of approximation (RMSEA) below 0.08, a standardized root mean square residual (SRMR) less than 0.08, and a comparative fit index (CFI) greater than 0.95 (Bentler 1990; Bowen and Guo 2011; Schreiber et al. 2006). These indices and suggested criteria were employed to test the fitness between the collected data and our structural model.

Weighting

In this study, when the two-stage stratified cluster sampling design was applied, the probabilities of individual students being sampled were unlikely to be equal because of the varied sizes of schools and classes. Such unequal probabilities may lead to under- or over-representation of the target population and to biased estimates from such a sample. Therefore, to better reflect the representativeness, sampling weights were calculated by using the inverse of individual students' probability of participation within the population stratum where the students belonged. The house weights were then estimated by linearly transforming the sampling weights so that the sum of the house weights was equal to the sample size. Finally, the house weights were applied and appropriated for the analyses that were sensitive to sample size (Rutkowski et al. 2010).

Controlling variables

As mentioned previously, the variables of computer experience and socioeconomic status were measured and controlled in this study in order to accurately determine the relationships between curiosity, engagement, and inquiry abilities. We used the strategy suggested by Spector and Brannick (2011) to eliminate the effects caused by computer experience and socioeconomic status through controlling the covariance of these two variables in the structural model. That is, we connected the two control variables to all the latent variables in the model. We then explained and reported the relationships among curiosity, dimensions of engagement, and inquiry abilities after the impact of the two control variables was considered and extracted.

Results

Descriptive and Correlational Results

The descriptive statistics of the data showed that the values of Kolmogorov–Smirnov of all indicators were significant ("Appendix A"). This result showed that the indicators in this study were normally distributed (Justel et al. 1997). Thus, the maximum likelihood (ML) was used to estimate the parameters of the structural equations (Bollen 1989).

Additionally, correlations between the indicators were examined. In general, most of these indicators were significantly related to inquiry abilities and related to each other (see "Appendix B"). However, some correlations of the indicators measuring the same latent variables were lower than their correlations with indicators of other latent variables. This implied that some indicators of the same variable (e.g., CE_5, CE_7) might not converge. In addition, we examined the reliability of the 37 indicators of curiosity and engagement ("Appendix C"). The initial Cronbach's α values ranged from 0.72 to 0.89. Although the values revealed that the reliability of our instrument was acceptable (Cronbach 1951), combining the reliability data with the correlational results, we removed nine items of engagement that indicated low convergence and low internal consistency (i.e., CE_5, CE_6, CE_7, BE_5, BE_6, EE_8, EE_9, SE_5, and SE_6). The Cronbach's α values were recalculated and ranged between 0.80 and 0.89 after nine indicators were deleted, which suggested a high internal consistency of the variables ("Appendix C"). Five indicators for inquiry-related curiosity and 23 for engagement (a total of 28 indicators) were used for the subsequent modeling analysis.

Measurement modeling

CFA was employed to confirm the validity of the five latent variables (i.e., inquiry-related curiosity, cognitive engagement, behavioral engagement, emotional engagement, and social engagement) based on the reliability and coherence of indicators within each latent variable. The CFA results presented the information about the convergent and discriminant patterns of the indicators and their corresponding latent variables in our measurement model. Table 3 shows the standardized factor loadings and residual variances of the indicators for each latent variable in both the initial and re-specified measurement models.

In our initial model, the factor loadings of the 28 indicators were significant. However, according to the suggested criteria, only 22 indicators were considered as reliable enough because their loadings were greater than 0.55. Six indicators of engagement (BE_4, BE_7, BE_8, EE_5, EE_6, and EE_7) with lower loadings were excluded from further analysis. Thus, the re-specified loadings of indicators revealed a satisfactory value to their corresponding variables.

Additionally, the fit indices derived from the CFA confirmed that the re-specified measurement model had an overall adequate model-data fit. Although the chi-square test of the re-specified measurement model showed a significant discrepancy between the re-specified model and the data ($\chi^2 = 820.38$, $df = 199$, $p < 0.001$), the ratio of chi-square per degree of freedom ($\frac{\chi^2}{df} = 4.12$) below 5 suggested that the re-specified measurement model fitted well to the data (Kelloway 1998; Lin and Tsai 2008). Moreover, the other indices indicated that our measurement model had an overall good fit with the data (RMSEA $= 0.072 < 0.080$, SRMR $= 0.049 < 0.08$, and CFI $= 0.98 > 0.95$).

After the fit indices were considered, the composite reliability (CR) and average variance extracted (AVE) indexes were estimated to examine the convergence of the five multi-indicator latent variables (i.e., four dimensions of engagement and inquiry-related curiosity). All of the CRs ranged between 0.76 and 0.93 and exceeded 0.60 (Table 3) so the five variables demonstrated sufficient internal consistency (Bagozzi and Yi 1988). In the final specified model, while the AVEs of the other four latent variables ranged from 0.51 to 0.72, the AVE of cognitive engagement was lower than 0.50. Yet, Fornell and Larcker (1981) indicated that a slightly lower AVE could be acceptable if the composite reliability of the latent variable was larger than 0.6 because, according to their simulations, the convergent validity of the variable was still adequate. As can be seen in Table 3, although the AVE of cognitive engagement was 0.46, the composite reliability was 0.83, much higher than the cut-off value of 0.6. Thus, together the AVEs and composite reliabilities revealed that the convergent validity of the latent variables in this study was acceptable.

Furthermore, the square roots of all five AVEs were calculated to estimate the discriminant validity as presented in Table 4. All but one of the square roots were greater than the correlations between each of the two latent variables. The correlation between cognitive engagement and social engagement reached 0.80, which was higher than 0.68 ($\sqrt{AVE}$ of CE) and 0.71 ($\sqrt{AVE}$ of SE). This implies a low discriminant validity and suggests that cognitive and social engagement had a strong inter-correlation with each other. We further

Table 4 Discriminant validity for the measurement model

Latent variable	Curi	CE	BE	EE	SE
Inquiry-related curiosity (Curi)	(0.85)[a]				
Cognitive engagement (CE)	0.64[b]	(0.68)			
Behavioral engagement (BE)	0.46	0.63	(0.96)		
Emotional engagement (EE)	0.50	0.62	0.67	(0.84)	
Social engagement (SE)	0.58	0.80	0.58	0.61	(0.71)

[a]Diagonal in parentheses: square roots of average variance extracted (AVE) from indicators

[b]Off-diagonal: correlations between constructs

examined the Heterotrait-Monotrait (HTMT) ratio to ensure the discriminant validity of the two constructs (Henseler et al. 2015). The HTMT of cognitive and social engagement was 0.807 in this study, lower than the suggested threshold of 0.85. These results confirmed the discriminant validity of all constructs.

Structural modeling

In the second phase of our analysis, the structural relationships proposed in Fig. 1 were examined and the fit indices were derived to support the fits between our structural model and the data. The ratio of chi-square per degree of freedom was less than 5.0 ($\chi^2 = 1035.82$, $df = 304$, $p < 0.001; = \frac{\chi^2}{df} 3.41$) and other fit indices indicated that our structural model had an overall good fit with the data (RMSEA = 0.064 < 0.080, SRMR = 0.057 < 0.08, and CFI = 0.97 > 0.95). The estimated model in Fig. 3 was thus adequate to describe the structural relationships between inquiry-related curiosity, the four dimensions of engagement, and inquiry abilities.

Interplay among the dimensions of engagement

The first research question was to examine the relationships among the four dimensions of engagement based on our hypothesized model. While Table 5 shows that the total effects of the behavioral, social, and emotional engagement on cognitive engagement were all significant, Fig. 3 reveals that the three dimensions affected cognitive engagement through different paths. The results showed that behavioral and social engagement were directly associated with cognitive engagement (effect = 0.20, $SE = 0.05$, $p < 0.001$ for behavioral engagement, and effect = 0.57, $SE = 0.08$, $p < 0.001$ for social

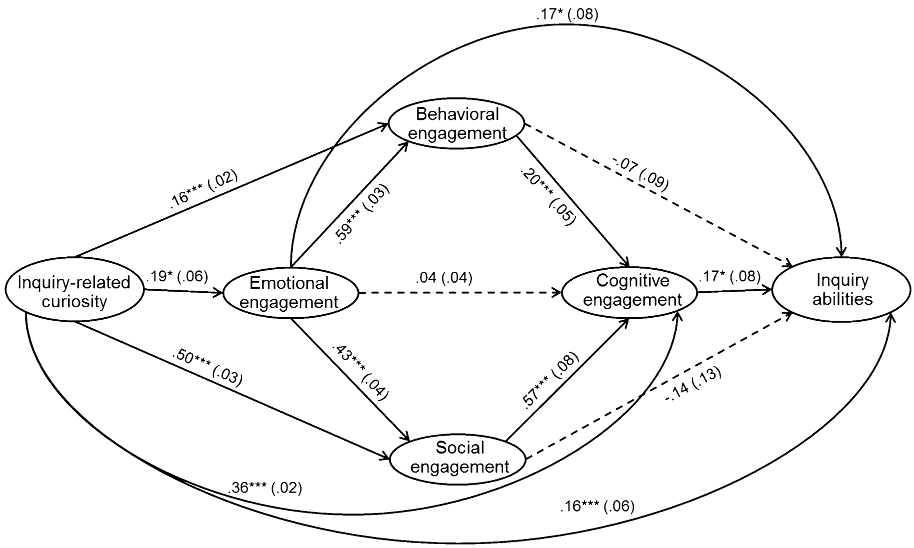

Fig. 3 The structural equation model of engagement in scientific inquiry. A bold line presents a significant association, a dotted arrow line presents an insignificant association, and the arrow presents the direction. The value presents the factor loading between the two variables. *$p < .05$. **$p < .01$. ***$p < .001$

Table 5 The standardized effects of behavioral, social, and emotional engagement on cognitive engagement

Exogenous variable	Direct effect (SE)	Indirect effect (SE)	Total effect (SE)
Behavioral engagement	0.20 (0.05)***	–[a]	0.20 (0.05)***
Social engagement	0.57 (0.08)***	–[a]	0.57 (0.08)***
Emotional engagement	0.04 (0.03)	0.36 (0.05)***	0.40 (0.04)***
Through behavioral engagement [b]	–	0.12 (0.03)***	–
Through social engagement [c]	–	0.24 (0.05)***	–
Difference[d]	–	0.12 (0.07)	–

[a]No indirect effect on cognitive engagement

[b]The pathway effect from emotional to cognitive engagement through behavioral engagement

[c]The pathway effect from emotional to cognitive engagement through social engagement

[d]The difference between the two pathway effects

***$p < .001$

engagement). On the other hand, emotional engagement did not directly link to cognitive engagement (direct effect = 0.04, $SE = 0.03$, *ns*) but affected it through the other two dimensions of engagement (indirect $\text{effect}_{(EE \rightarrow BE \rightarrow CE)} = 0.12$, $SE = 0.03$, $p < 0.001$ and indirect $\text{effect}_{(EE \rightarrow SE \rightarrow CE)} = 0.24$, $SE = 0.05$, $p < 0.001$). The total indirect effects of emotional engagement on cognitive engagement was significant (total indirect effect = 0.36, $SE = 0.05$, $p < 0.001$). Furthermore, a comparison between the two indirect effects of emotional engagement revealed no significant difference between the two pathway effects (difference = 0.12, $SE = 0.07$, *ns*). This result indicated that behavioral and social engagement had similar power to mediate the effect between emotional and cognitive engagement. Overall, the results partially supported our hypothesis by showing that emotional engagement could predict behavioral and social engagement, but emotional engagement could only indirectly affect cognitive engagement.

Effects of the four dimensions of engagement on inquiry abilities

To answer the second research question, the effects of the four dimensions of engagement on inquiry abilities were investigated. As shown in Table 6, our hypothesis about the positive influence of the four dimensions of engagement on inquiry abilities was not fully confirmed. Only the total effects of cognitive and emotional engagement were significant (total effect = 0.17, $SE = 0.08$, $p < 0.05$ for cognitive engagement and total effect = 0.14, $SE = 0.04$, $p < 0.001$ for emotional engagement). Among the four dimensions, cognitive and emotional engagement played a relatively important role in enhancing students' inquiry abilities, and directly affected their inquiry abilities ($\text{effect}_{(CE \rightarrow ability)} = 0.17$, $SE = 0.08$, $p < 0.05$ and $\text{effect}_{(EE \rightarrow ability)} = 0.17$, $SE = 0.08$, $p < 0.05$). Both of the two variables had the same explanation of variance ($R^2 = 2.89\%$) to predict students' inquiry abilities, separately. Yet, a close examination of the effects revealed that although the total effects of behavioral and social engagement on inquiry abilities were not significant, they indirectly affected students' inquiry

Table 6 The standardized effects on inquiry abilities

Exogenous variable	Direct effect (SE)	Indirect effect (SE)	Total effect (SE)
Cognitive engagement (CE)	0.17 (0.08)*	–[a]	0.17 (0.08)*
Behavioral engagement (BE)	− 0.07 (0.09)	0.03 (0.02)*	− 0.04 (0.08)
Social engagement (SE)	− 0.14 (0.13)	0.09 (0.05)*	− 0.05 (0.16)
Emotional engagement (EE)	0.17 (0.08)*	− 0.03 (0.02)	0.14 (0.04)***
EE → CE → Ability	–	0.01 (0.01)	–
EE → BE → Ability	–	− 0.04 (0.05)	–
EE → BE → CE → Ability	–	0.02 (0.01)*	–
EE → SE → Ability	–	− 0.06 (0.06)	–
EE → SE → CE → Ability	–	0.04 (0.02)*	–
Inquiry-related curiosity	0.16 (0.03)***	0.06 (0.02)**	0.22 (0.07)**
Curi → EE → Ability	–	0.03 (0.02)	–
Curi → CE → Ability	–	0.06 (0.03)*	–
Curi → EE → BE → CE → Ability	–	0.00 (0.00)*	–
Curi → EE → BE → Ability	–	− 0.01 (0.01)	–
Curi → EE → SE → CE → Ability	–	0.01 (0.00)*	–
Curi → EE → SE → Ability	–	− 0.01 (0.01)	–
Curi → EE → CE → Ability	–	0.00 (0.00)	–
Curi → BE → Ability	–	− 0.01 (0.01)	–
Curi → BE → CE → Ability	–	0.01 (0.00)*	–
Curi → SE → Ability	–	− 0.07 (0.07)	–
Curi → SE → CE → Ability	–	0.05 (0.02)*	–

Ability inquiry ability, *BE* behavior engagement, *CE* cognitive engagement, *Curi* inquiry-related curiosity, *EE* emotional engagement, *SE* social engagement, *SE* standard error

[a]No indirect effect on inquiry abilities

$*p<.05$. $**p<.01$. $***p<.001$

abilities, and their significant indirect influence on inquiry abilities was completely mediated by cognitive engagement ($\text{effect}_{(BE \rightarrow CE \rightarrow ability)} = 0.03$, $SE = 0.02$, $p < 0.05$ and $\text{effect}_{(SE \rightarrow CE \rightarrow ability)} = 0.09$, $SE = 0.05$, $p < 0.05$).

Inquiry-related curiosity affecting the dimensions of engagement

The third research question focused on the associations between students' inquiry-related curiosity and the dimensions of engagement. Figure 3 and Table 7 illustrate that the total effects of curiosity on the four dimensions of engagement were all significant (total effect (Curiosity → BE) = 0.27, $SE = 0.05$, $p < 0.001$; total effect (Curiosity → SE) = 0.58, $SE = 0.04$, $p < 0.001$; total effect (Curiosity → EE) = 0.19, $SE = 0.06$, $p < 0.001$; total effect (Curiosity → CE) = 0.75, $SE = 0.04$, $p < 0.001$). The results supported our hypothesis that inquiry-related curiosity had a positive impact on the four dimensions of engagement. Especially, the total effect of curiosity on cognitive engagement was the strongest. The

Table 7 The standardized effects of inquiry-related curiosity on the dimensions of engagement

Exogenous variable	Direct effect (SE)	Indirect effect (SE)	Total effect (SE)
On behavioral engagement	0.16 (0.02)***	0.11 (0.03)***	0.27 (0.05)***
On social engagement	0.50 (0.03)***	0.08 (0.02)***	0.58 (0.04)***
On emotional engagement	0.19 (0.06)*	–[a]	0.19 (0.06)***
On cognitive engagement	0.36 (0.02)***	0.39 (0.02)***	0.75 (0.04)***
Curi → BE → CE	–	0.03 (.01)***	–
Curi → SE → CE	–	0.29 (.04)***	–
Curi → EE → CE	–	0.01 (.01)	–
Curi → EE → BE → CE	–	0.02 (.01)*	–
Curi → EE → SE → CE	–	0.05 (.01)***	–

BE behavior engagement, *CE* cognitive engagement, *Curi* inquiry-related curiosity, *EE* emotional engagement, *SE* social engagement, *SE* standard error

[a]No indirect effect on emotional engagement

**p* < .05. ****p* < .001

coefficient indicated that inquiry-related curiosity could explain the 56.25% covariance of cognitive engagement with the mediation of emotional, behavioral, and social engagement.

Additionally, Table 7 shows that inquiry-related curiosity significantly affected all four dimensions of engagement directly and that its direct effect on social engagement was particularly strong. Moreover, curiosity was indirectly, significantly associated with cognitive, behavioral, and social engagement. As shown in Fig. 3, its indirect effects on behavioral and social engagement were through emotional engagement ($\text{effect}_{(\text{Curiosity}\rightarrow\text{EE}\rightarrow\text{BE})} = 0.11$, $SE = 0.03$, $p < 0.01$ and $\text{effect}_{(\text{Curiosity}\rightarrow\text{EE}\rightarrow\text{SE})} = 0.08$, $SE = 0.02$, $p < 0.01$). On the other hand, the indirect influence on cognitive engagement was through multiple pathways (Table 7). Among them, the strongest indirect effect was through social engagement ($\text{effect}_{(\text{Curiosity}\rightarrow\text{SE}\rightarrow\text{CE})} = 0.29$, $SE = 0.04$, $p < 0.01$). Together these results suggested that curiosity was an important factor in promoting engagement in science activities.

Direct and indirect effects of curiosity on students' inquiry abilities

The fourth research question addressed the impact of curiosity on students' inquiry abilities. The results showed the significant direct, indirect, and total effects of curiosity on inquiry abilities (direct effect = 0.16. $SE = 0.05$, $p < 0.001$; indirect effect = 0.06. $SE = 0.02$, $p < 0.01$; total effect = 0.22. $SE = 0.02$, $p < 0.01$, for details see in Table 6). Our hypothesis on the mediating role of engagement was not fully confirmed; as can be seen in Table 2, the results suggested that while some pathways mediated by engagement were significant, others were not. Among them, the pathways involving cognitive engagement were all significant except for one (Curi → EE → CE → Ability). Overall, Fig. 3 illustrates that the associations between curiosity and the four dimensions of engagement played a partially mediated role in the relationship between students' curiosity and inquiry abilities.

Discussion and implications

Interplay among the dimensions of engagement

Previous studies have suggested inter-relationships among some dimensions of engagement (e.g., Gunuc and Kuzu 2015; Li and Lerner 2013) and this study extends current knowledge by demonstrating the internal dynamics of engagement in a disciplinary context. While this study echoes the findings from the previous study by confirming the predictive links from emotional engagement to behavioral and social engagement, our results do not support a direct effect of emotional engagement on cognitive engagement, but suggest an indirect influence through behavioral and social engagement. That is, during scientific inquiry, students' positive perceptions and feelings about inquiry activities may not determine their use of cognitive strategies and pursuit of achievement; rather, students' emotional engagement could trigger their attentiveness and on-task behaviors and enhance their collaboration with others, which in turn would increase their cognitive efforts invested in scientific inquiry.

These results related to the internal dynamics of engagement can provide useful information for teachers and designers of science learning environments to facilitate a certain dimension of engagement. Teachers could design interesting and meaningful science activities to enhance students' perceptions and values of learning environments. Such high emotional engagement could lead to students' attentiveness and investment of time, and may further support their cognitive engagement. Additionally, encouraging students to work in groups and increasing their interactions with peers could promote social engagement that may energize them to engage cognitively and to accomplish learning tasks.

Enhancing inquiry abilities through engagement

Consistent with the findings from previous research, this study supports the positive association between cognitive engagement and inquiry abilities. The significant direct effect of emotional engagement on inquiry abilities suggests that students' positive feelings and perceptions of the inquiry activities in science laboratory classes may lead to better performances in asking questions, analyzing data, doing experiments, and making explanations during the scientific inquiry.

However, this result is not consistent with the finding of Wu and Hang (2007), which indicated that the level of emotional engagement did not affect students' achievement. There may be several reasons for the different results of the two studies. First, to examine whether emotional engagement affected students' learning achievement, Wu and Hang (2007) divided students into three groups based on their scores on the questionnaire of emotional engagement, and conducted a 2 (instructional approach) $\times$ 3 (emotional engagement level) two-way ANOVA test on the achievement post-test. That is, Wu and Hang (2007) categorized the continuous variable of emotional engagement to compare group differences; however, according to Maxwell and Delaney (1993) and Altman and Royston (2006), such a grouping method could cause a reduction in statistical power and lead to false statistical significance. This statistical weakness caused by grouping was resolved in this study, which may have allowed a more accurate estimation of the effect from emotional engagement. Additionally, the statistical methods (2-way ANOVA) used in Wu and Hang (2007) did not consider effects from other dimensions of engagement and curiosity. On the other hand, this study controlled possible confounding variables and used SEM to identify direct and indirect effects of latent variables. The statistical methods used in this study may be more appropriate to estimate the effects of engagement on learning. Thirdly, the learning contexts were different in the two studies. While the Wu and Hang (2007) study was set in technology-enhanced science learning environments, this study focused on science laboratory classes. The nature of engagement may be different in these two types of learning contexts, possibly leading to diverse effects. Finally, the two studies measured different learning outcomes. The learning outcome measured in this study was inquiry abilities, whereas the learning achievement in Wu and Hang (2007) was students' conceptual knowledge. Different learning outcomes may have different associations to engagement. By employing rigorous statistical methods, this study shows an evident association between emotional engagement and inquiry abilities, and emotional engagement is significant in promoting science learning when students participate in inquiry-based, hands-on activities to develop skills or abilities. An important practical implication may be that an encouraging learning environment where students' positive perceptions about science inquiry are facilitated may support students' development of inquiry abilities.

Additionally, the question of whether social engagement affects students' learning achievement has been relatively under-investigated in early research. This study indicates that although social engagement did not directly affect students' inquiry abilities, its indirect effect was significant and completely mediated by cognitive engagement. These results support the benefits of collaborative learning. The results also imply that students' interactions with peers could enhance their use of cognitive strategies and increase intellectual investment in inquiry activities, which may then promote better performances in inquiry.

Inquiry-related curiosity affecting engagement and inquiry abilities

Previous studies have indicated that curiosity can stimulate students to explore why and how a phenomenon occurs, and suggested that such a mechanism-searching process could facilitate their learning in science (Chin and Osborne 2008; Luce and Hsi 2015). This study advances current knowledge of how curiosity affects engagement and inquiry abilities by differentiating the pathways from students' curiosity to the four dimensions of engagement and inquiry abilities. First, the results of this study underscore the critical role of inquiry-related curiosity in promoting students' engagement in all four dimensions. The significant direct and indirect relationships imply that students' curiosity about scientific inquiry could evoke students' positive emotions, sustain their on-task behaviors, encourage interactions with peers, and initiate the use of cognitive strategies. Through different dimensions of engagement, students' inquiry abilities could be facilitated. Additionally, the direct effect on inquiry abilities indicates that inquiry-related curiosity may activate the development of inquiry abilities because students with a higher level of curiosity could be more willing to formulate scientific questions, conduct experiments and generate scientific explanations.

If curiosity is such an important indicator in scientific inquiry, how can teachers and educators promote students' curiosity in scientific inquiry? Increasing the frequency of students' science laboratory classes and their learning opportunities in informal science-designed settings (e.g., science camps, museums, and environmental centers) may be beneficial (Wu et al. 2018). Furthermore, teachers could identify the topics or questions that intrigue students, and design inquiry tasks around those topics (Krajcik and Czerniak 2007). Allowing students to explore questions they are curious about could sustain their engagement in scientific inquiry and could subsequently promote the development of their inquiry abilities.

Limitations of the study

Although this study advances understanding of the relationships between curiosity, engagement, and inquiry abilities, the results of this study were subject to the following limitations. The first limitation is regarding our assessment framework. The inquiry abilities covered in this study were limited to four inquiry phases: asking scientific questions, planning experiments, analyzing data, and formulating scientific explanations. Other inquiry abilities, such as modeling, argumentation, and communication, were not examined in this study. In addition, this study viewed the inquiry abilities as a unidimensional construct, and unidimensional Rasch models were used to estimate students' inquiry abilities. This study did not estimate plausible values for each of the four abilities because our previous study (Kuo et al. 2015) supported a unidimensional structure and showed that "the correlations among the four inquiry components were higher than the person separation reliabilities of the four components." Future studies may thus consider including more inquiry abilities in their assessment frameworks and when more abilities are evaluated, the structure of the constructs may change and different inquiry abilities may be viewed as having multiple dimensions.

A second limitation is from our choice of factors that affect students' inquiry abilities. Although the four dimensions of engagement significantly affected inquiry abilities directly or indirectly (Table 6), the effect sizes were small and the covariance of inquiry abilities that could be explained by engagement was low. The results suggest that other motivational and cognitive variables may be associated with inquiry abilities. For example, Nehring et al. (2015) used the hierarchical regression analysis to investigate how multiple student charateritics contribute to inquiry skills in chemistry. They found that the cognitive variables (e.g., conceptual knowledge in chemistry, intelligence, and perceived cognitive load) predicted 47% of the inquiry skills while the motivational variables (e.g., interest in scientific investigation and self-concept in chemistry) explained these skills up to 25%. Future research may draw upon different theoretical perspectives, take more variables into consideration, and establish a more comprehensive model to explain the development of inquiry abilities.

Thirdly, in the engagement questionnaire, instead of using the term "inquiry," we described the learning context as a "science laboratory" (Table 1). The research literature in science education has made a clear distinction between science laboratory and inquiry (e.g., Hofstein and Lunetta 2004); however, previous research also revealed the variety of meanings associated with the term inquiry (e.g., Abd-El-Khalick et al. 2004). In Taiwan, the term inquiry was not formally introduced to teachers and students until the new science curriculum guidelines were launched in 2019. When this study was conducted, considering students' unfamiliarity with "inquiry," we decided to carefully describe the activities in the items and use common terms in science classes in Taiwan, such as experiment and laboratory, to refer to inquiry. This was a difficult research decision because while the choice of inappropriate terms may undermine the validity of the questionnaire, participants' misinterpretations and lack of knowledge of the terms may reduce the quality of the data (Krosnick and Presser 2010). This limitation could be resolved in our future study as new inquiry-based science curricula are implemented and Taiwanese teachers and students become more familiar with science inquiry.

Conclusion

Researchers and educators in science education advocate a focus on scientific inquiry, and promoting students' inquiry abilities has been identified as a key objective of science education. This study sheds light on how to foster students' inquiry abilities by enhancing different dimensions of engagement and promoting students' curiosity. This study constructed a model of engagement in scientific inquiry and explored how four dimensions of engagement (i.e., cognitive, behavioral, emotional, and social) were driven by inquiry-related curiosity and affected students' inquiry abilities. The results showed that students' curiosity was associated with their inquiry abilities, and such an association was partially mediated by the four dimensions of engagement in science laboratory classes. Moreover, among the four dimensions of engagement, only cognitive and emotional engagement had significant total effects on students' inquiry abilities, and the influence of behavioral and social engagement on inquiry abilities was completely mediated by cognitive engagement. These results suggest a critical role played by emotional engagement, cognitive engagement, and curiosity in science inquiry. Therefore, to develop students' inquiry abilities, teachers and educators should create an encouraging and collaborative learning environment where students could follow their curiosity, have positive perceptions of science inquiry, and be willing to make cognitive efforts in inquiry activities.

Acknowledgements This study was finally supported by the Ministry of Science and Technology in Taiwan under MOST 103–2511-S-003–038-MY4, MOST 106–2511-S-003–046-MY3, and the "Institute for Research Excellence in Learning Sciences" of National Taiwan Normal University from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education in Taiwan.

Appendix A

Descriptive statistics of the indicators.

Indicators	Pv_1	Curi1	Curi2	Curi3	Curi4	Curi5	CE_1	CE_2	CE_3	CE_4
N	605	605	605	605	605	605	605	605	605	605
Kolmogorov–Smirnova	0.056***	0.365***	0.346***	0.313***	0.331***	0.317***	0.245***	0.235***	0.236***	0.287***
Mean	0.247	3.14	3.09	3.15	3.21	3.06	2.82	2.74	2.67	2.99
Std. Deviation	0.663	0.601	0.628	0.656	0.628	0.678	0.755	0.799	0.793	0.747
Variance	0.440	0.362	0.395	0.43	0.395	0.46	0.571	0.639	0.629	0.558
Minimum	− 1.891	1	1	1	1	1	1	1	1	1
Maximum	2.150	4	4	4	4	4	4	4	4	4

Indicators	CE_5	CE_6	CE_7	CE_8	CE_9	BE_1	BE_2	BE_3	BE_4
N	605	605	605	605	605	605	605	605	605
Kolmogorov–Smirnova	0.300***	0.293***	0.240***	0.251***	0.224***	0.269***	0.271***	0.290***	0.231***
Mean	2.92	2.9	2.47	2.58	2.65	3.04	3.07	2.94	2.67
Std. Deviation	0.777	0.775	0.815	0.827	0.836	0.722	0.716	0.724	0.793
Variance	0.604	0.6	0.664	0.685	0.699	0.522	0.512	0.525	0.629
Minimum	1	1	1	1	1	1	1	1	1
Maximum	4	4	4	4	4	4	4	4	4

Indicators	BE_5	BE_6	BE_7	BE_8	EE_1	EE_2	EE_3	EE_4	EE_5	EE_6
N	605	605	605	605	605	605	605	605	605	605
Kolmogorov–Smirnova	0.319***	0.492***	0.271***	0.394***	0.196***	0.230***	0.260***	0.262***	0.285***	0.410***
Mean	2.24	3.76	3.29	3.55	2.82	2.87	2.96	2.88	3.33	3.56

Indicators	BE_5	BE_6	BE_7	BE_8	EE_1	EE_2	EE_3	EE_4	EE_5	EE_6
Std. Deviation	0.748	0.558	0.68	0.663	0.906	0.842	0.797	0.817	0.734	0.713
Variance	0.559	0.312	0.462	0.44	0.821	0.709	0.636	0.668	0.539	0.509
Minimum	1	1	1	1	1	1	1	1	1	1
Maximum	4	4	4	4	4	4	4	4	4	4

Indicators	EE_7	EE_8	EE_9	SE_1	SE_2	SE_3	SE_4	SE_5	SE_6
N	605	605	605	605	605	605	605	605	605
Kolmogorov–Smirnova	0.438***	0.458***	0.437***	0.257***	0.257***	0.297***	0.290***	0.309***	0.366***
Mean	3.61	3.69	3.63	2.55	2.6	2.94	2.78	3.33	3.49
Std. Deviation	0.678	0.598	0.626	0.765	0.788	0.733	0.748	0.804	0.729
Variance	0.46	0.358	0.392	0.585	0.621	0.538	0.559	0.647	0.532
Minimum	1	1	1	1	1	1	1	1	1
Maximum	4	4	4	4	4	4	4	4	4

[a]The value of Kolmogorov–Smirnov represents the tests of normality

*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$

Appendix B

Correlation matrix of the indicators (N = 605).

Indicators	pv_1	Curi1	Curi2	Curi3	Curi4	Curi5	CE_1	CE_2	CE_3	CE_4
pv_1	1									
Curi1	0.170**	1								
Curi2	0.132**	0.637**	1							
Curi3	0.214**	0.574**	0.663**	1						
Curi4	0.219**	0.630**	0.616**	0.698**	1					
Curi5	0.160**	0.576**	0.571**	0.620**	0.662**	1				
CE_1	0.109**	0.293**	0.263**	0.269**	0.330**	0.344**	1			
CE_2	0.078	0.356**	0.322**	0.327**	0.369**	0.337**	0.431**	1		
CE_3	0.122**	0.332**	0.288**	0.294**	0.349**	0.348**	0.346**	0.618**	1	
CE_4	0.210**	0.390**	0.383**	0.364**	0.445**	0.396**	0.508**	0.521**	0.526**	1
CE_5	0.05	0.005	0.046	0.119**	0.013	0.016	0.07	0.117**	0.019	0.079
CE_6	0.099*	0.089*	0.104*	0.119**	0.082*	0.103*	0.062	0.112**	0.017	0.130**
CE_7	− 0.06	− 0.021	− 0.036	− 0.036	− 0.062	− 0.042	− 0.099*	0.120**	0.092*	0.045
CE_8	0.076	0.268**	0.275**	0.265**	0.272**	0.275**	0.391**	0.446**	0.433**	0.399**
CE_9	0.142**	0.312**	0.302**	0.366**	0.324**	0.338**	0.304**	0.446**	0.508**	0.467**
BE_1	0.135**	0.247**	0.288**	0.289**	0.308**	0.303**	0.509**	0.335**	0.316**	0.455**
BE_2	0.157**	0.305**	0.344**	0.359**	0.385**	0.347**	0.487**	0.363**	0.357**	0.493**
BE_3	0.165**	0.354**	0.360**	0.395**	0.412**	0.374**	0.489**	0.495**	0.402**	0.574**
BE_4	0.055	0.239**	0.198**	0.227**	0.219**	0.231**	0.357**	0.247**	0.234**	0.289**
BE_5	0.006	0.288**	0.283**	0.287**	0.281**	0.295**	0.251**	0.352**	0.395**	0.347**
BE_6	0.245**	0.034	0.041	0.100*	0.063	0.064	0.154**	0.025	0.011	0.158**
BE_7	0.128**	0.031	0.079	0.130**	0.089*	0.056	0.192**	0.072	0.02	0.125**
BE_8	0.188**	0.158**	0.150**	0.187**	0.174**	0.129**	0.171**	0.153**	0.077	0.240**
EE_1	0.210**	0.250**	0.298**	0.332**	0.321**	0.303**	0.365**	0.312**	0.239**	0.348**
EE_2	0.183**	0.267**	0.370**	0.382**	0.354**	0.336**	0.407**	0.406**	0.350**	0.441**
EE_3	0.213**	0.344**	0.424**	0.452**	0.411**	0.418**	0.392**	0.456**	0.397**	0.480**
EE_4	0.174**	0.244**	0.278**	0.287**	0.326**	0.264**	0.338**	0.346**	0.279**	0.386**
EE_5	0.063	0.02	0.124**	0.143**	0.076	0.110**	0.136**	0.097*	0.025	0.135**
EE_6	0.129**	0.033	0.124**	0.165**	0.123**	0.126**	0.133**	0.104*	0.03	0.195**
EE_7	0.203**	0.113**	0.189**	0.237**	0.149**	0.202**	0.207**	0.145**	0.152**	0.265**
EE_8	0.131**	0.047	0.086*	0.100*	0.082*	0.066	0.085*	0.018	0.023	0.131**
EE_9	0.160**	0.03	0.073	0.068	0.091*	0.036	0.04	− 0.01	0.012	0.086*
SE_1	0.075	0.352**	0.354**	0.314**	0.365**	0.370**	0.334**	0.499**	0.500**	0.445**
SE_2	0.096*	0.278**	0.368**	0.357**	0.343**	0.379**	0.320**	0.392**	0.410**	0.442**
SE_3	0.171**	0.241**	0.231**	0.291**	0.355**	0.304**	0.300**	0.295**	0.332**	0.377**
SE_4	0.139**	0.284**	0.267**	0.294**	0.320**	0.310**	0.396**	0.420**	0.397**	0.453**
SE_5	0.05	0.037	0.096*	0.082*	0.054	0.049	0.085*	0.021	0.064	0.071
SE_6	0.126**	0.048	0.084*	0.129**	0.077	0.099*	0.077	0.088*	0.109**	0.108**

Indicators	CE_5	CE_6	CE_7	CE_8	CE_9	BE_1	BE_2	BE_3	BE_4
pv_1									
Curi1									
Curi2									
Curi3									
Curi4									
Curi5									
CE_1									
CE_2									
CE_3									
CE_4									
CE_5	1								
CE_6	0.610**	1							
CE_7	0.309**	0.393**	1						
CE_8	0.109**	0.068	0.049	1					
CE_9	0.069	0.086*	0.080*	0.494**	1				
BE_1	0.109**	0.143**	− 0.080*	0.264**	0.312**	1			
BE_2	0.087*	0.150**	− 0.011	0.288**	0.352**	0.866**	1		
BE_3	0.150**	0.264**	0.061	0.360**	0.409**	0.646**	0.675**	1	
BE_4	0.108**	0.100*	− 0.131**	0.236**	0.201**	0.504**	0.473**	0.465**	1
BE_5	0.009	0.057	0.005	0.382**	0.439**	0.303**	0.342**	0.380**	0.353**
BE_6	0.274**	0.287**	0.138**	0.018	0.018	0.165**	0.201**	0.150**	0.019
BE_7	0.354**	0.401**	0.159**	0.058	0.071	0.292**	0.273**	0.219**	0.174**
BE_8	0.354**	0.441**	0.235**	0.144**	0.127**	0.176**	0.230**	0.322**	0.077
EE_1	0.094*	0.212**	0.094*	0.245**	0.265**	0.447**	0.462**	0.479**	0.218**
EE_2	0.160**	0.235**	0.102*	0.328**	0.370**	0.558**	0.570**	0.579**	0.298**
EE_3	0.123**	0.178**	0.099*	0.380**	0.438**	0.520**	0.556**	0.581**	0.277**
EE_4	0.088*	0.174**	0.077	0.234**	0.317**	0.450**	0.449**	0.503**	0.259**
EE_5	0.330**	0.354**	0.165**	0.076	0.049	0.268**	0.277**	0.266**	0.133**
EE_6	0.264**	0.324**	0.154**	0.098*	0.066	0.247**	0.252**	0.284**	0.115**
EE_7	0.293**	0.348**	0.151**	0.144**	0.156**	0.323**	0.333**	0.345**	0.195**
EE_8	0.182**	0.248**	0.142**	0.015	0.031	0.172**	0.207**	0.188**	0.096*
EE_9	0.147**	0.239**	0.134**	− 0.06	− 0.01	0.125**	0.176**	0.117**	0.025
SE_1	0.045	0.092*	0.068	0.381**	0.519**	0.380**	0.384**	0.476**	0.255**
SE_2	0.056	0.109**	0.041	0.388**	0.451**	0.372**	0.402**	0.467**	0.373**
SE_3	− 0.029	− 0.005	− 0.082*	0.282**	0.332**	0.351**	0.361**	0.330**	0.276**
SE_4	0.065	0.092*	0.008	0.316**	0.419**	0.444**	0.429**	0.481**	0.296**
SE_5	0.220**	0.247**	0.203**	0.065	0.05	0.150**	0.149**	0.104*	0.171**
SE_6	0.183**	0.236**	0.188**	0.085*	0.200**	0.130**	0.157**	0.107**	0.089*

Indicators	BE_5	BE_6	BE_7	BE_8	EE_1	EE_2	EE_3	EE_4	EE_5	EE_6
pv_1										
Curi1										
Curi2										
Curi3										

Indicators	BE_5	BE_6	BE_7	BE_8	EE_1	EE_2	EE_3	EE_4	EE_5	EE_6
Curi4										
Curi5										
CE_1										
CE_2										
CE_3										
CE_4										
CE_5										
CE_6										
CE_7										
CE_8										
CE_9										
BE_1										
BE_2										
BE_3										
BE_4										
BE_5	1									
BE_6	− 0.121**	1								
BE_7	0.061	0.508**	1							
BE_8	0.008	0.577**	0.522**	1						
EE_1	0.287**	0.204**	0.210**	0.210**	1					
EE_2	0.361**	0.189**	0.236**	0.230**	0.778**	1				
EE_3	0.414**	0.171**	0.218**	0.222**	0.682**	0.835**	1			
EE_4	0.304**	0.184**	0.236**	0.195**	0.633**	0.634**	0.628**	1		
EE_5	0.101*	0.350**	0.424**	0.381**	0.371**	0.377**	0.315**	0.300**	1	
EE_6	0.055	0.448**	0.435**	0.430**	0.465**	0.406**	0.323**	0.350**	0.677**	1
EE_7	0.143**	0.449**	0.444**	0.489**	0.386**	0.405**	0.374**	0.296**	0.546**	0.677**
EE_8	− 0.037	0.430**	0.376**	0.398**	0.216**	0.206**	0.171**	0.249**	0.457**	0.525**
EE_9	− 0.025	0.340**	0.323**	0.361**	0.128**	0.136**	0.106**	0.177**	0.328**	0.355**
SE_1	0.447**	− 0.01	0.045	0.090*	0.334**	0.413**	0.450**	0.320**	0.074	0.091*
SE_2	0.464**	0.058	0.118**	0.135**	0.354**	0.461**	0.464**	0.374**	0.109**	0.138**
SE_3	0.292**	0.110**	0.069	0.076	0.328**	0.333**	0.350**	0.364**	0.114**	0.132**
SE_4	0.375**	0.084*	0.107**	0.092*	0.407**	0.419**	0.421**	0.431**	0.162**	0.199**
SE_5	0.082*	0.293**	0.299**	0.306**	0.129**	0.198**	0.117**	0.101*	0.285**	0.316**
SE_6	0.082*	0.396**	0.337**	0.335**	0.160**	0.201**	0.171**	0.144**	0.261**	0.316**

Indicators	EE_7	EE_8	EE_9	SE_1	SE_2	SE_3	SE_4	SE_5	SE_6
pv_1									
Curi1									
Curi2									
Curi3									
Curi4									
Curi5									
CE_1									
CE_2									

Indicators	EE_7	EE_8	EE_9	SE_1	SE_2	SE_3	SE_4	SE_5	SE_6
CE_3									
CE_4									
CE_5									
CE_6									
CE_7									
CE_8									
CE_9									
BE_1									
BE_2									
BE_3									
BE_4									
BE_5									
BE_6									
BE_7									
BE_8									
EE_1									
EE_2									
EE_3									
EE_4									
EE_5									
EE_6									
EE_7	1								
EE_8	0.496**	1							
EE_9	0.397**	0.574**	1						
SE_1	0.156**	− 0.03	− 0.022	1					
SE_2	0.235**	0.014	0.017	0.663**	1				
SE_3	0.155**	0.024	0.048	0.406**	0.474**	1			
SE_4	0.223**	0.061	0.079	0.490**	0.501**	0.549**	1		
SE_5	0.396**	0.265**	0.178**	0.063	0.246**	0.068	0.097*	1	
SE_6	0.349**	0.394**	0.297**	0.058	0.163**	0.139**	0.163**	0.474**	1

PV_1 represents the first of the plausible values to which the indicator of students' inquiry abilities refers. Curi1 to Curi5 represent the indicators of the inquiry-related curiosity. CE_1 to CE_9 represent the indicators of the students' cognitive engagement. BE_1 to BE_8 represent the indicators of the students' behavioral engagement. EE_1 to EE_9 represent the indicators of the students' emotional engagement. SE_1 to SE_6 represent the indicators of the students' social engagement.

$*p<0.05$. $**p<0.01$. $***p<0.001$

Appendix C

Validation of the 37 indicators (N = 605).

Variable/indicators	Cronbach's α	Cronbach's α if Indicator Deleted
Curi	0.89(5)a	–
Curi_1	0.707	–
Curi_2	0.732	–
Curi_3	0.757	–
Curi_4	0.777	–
Curi_5	0.712	–
CE	0.77 (9)a	0.83 (6)[a]
CE_1	–	0.75
CE_2	–	0.72
CE_3	–	0.73
CE_4	–	0.73
CE_5[b]	–	0.77
CE_6[b]	–	0.76
CE_7[b]	–	0.78
CE_8	–	0.73
CE_9	–	0.73
BE	0.80 (8)a	0.80 (6)[a]
BE_1	–	0.74
BE_2	–	0.74
BE_3	–	0.75
BE_4	–	0.78
BE_5[b]	–	0.81
BE_6[b]	–	0.80
BE_7	–	0.79
BE_8	–	0.79
EE	0.87 (9)a	0.88 (7)[a]
EE_1	–	0.84
EE_2	–	0.84
EE_3	–	0.85
EE_4	–	0.85
EE_5	–	0.85
EE_6	–	0.85
EE_7	–	0.85
EE_8[b]	–	0.86
EE_9[b]	–	0.87
SE	0.72 (6)a	0.81 (4)[a]
SE_1	–	0.66
SE_2	–	0.62
SE_3	–	0.68
SE_4	–	0.66
SE_5[b]	–	0.74
SE_6[b]	–	0.73

[a]The numbers in the parentheses show the numbers of indicators for each variable
[b]The indicators were deleted for better reliabilities

References

Abd-El-Khalick, F., Boujaoude, S., Duschl, R., Lederman, N. G., Mamlok-Naaman, R., Hofstein, A., et al. (2004). Inquiry in science education: International perspectives. *Science Education, 88*, 394–419.

Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *British Medical Journal, 332*(7549), 1080. https://doi.org/10.1136/bmj.332.7549.1080.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*(3), 411–423.

Apedoe, X. S. (2008). Engaging students in inquiry: Tales from an undergraduate geology laboratory-based course. *Science Education, 92*(4), 631–663. https://doi.org/10.1002/sce.20254.

Archambault, I., Janosz, M., Fallu, J. S., & Pagani, L. S. (2009). Student engagement and its relationship with early high school dropout. *Journal of Adolescence, 32*(3), 651–670. https://doi.org/10.1016/j.adolescence.2008.06.007.

Astin, A. W. (1984). Student involvement: A developmental theory for higher education. *Journal of college student personnel, 25*(4), 297–308.

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science, 16*(1), 74–94. https://doi.org/10.1007/bf02723327.

Barrow, L. H. (2006). A brief history of inquiry: From dewey to standards. *Journal of Science Teacher Education, 17*(3), 265–278. https://doi.org/10.1007/s10972-006-9008-5.

Bell, D. (2002). Making science inclusive: Providing effective learning opportunities for children with learning difficulties. *Support for Learning, 17*(4), 156–161. https://doi.org/10.1111/1467-9604.00258.

Bentler, P. (1990). Comparative fit indices in structural models. *Psychological Bulletin, 107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Bowen, N. K., & Guo, S. (2011). Evaluating and improving CFA and general structural models. In N. K. Bowen & S. Guo (Eds.), Structural equation modeling: Oxford Scholarship Online. Retrieved from https://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195367621.001.0001/acprof-9780195367621-chapter-0006. doi:10.1093/acprof:oso/9780195367621.003.0006

Callanan, M. A., & Jipson, J. (2001). Explanatory conversations and young children's developing scientific literacy. In K. Crowley, C. D. Schunn, & T. Okada (Eds.), *Designing for science: Implications from everyday, classroom, and professional settings* (pp. 21–49). Mahwah, NJ: Lawrence Erlbaum Associations.

Callanan, M. A., & Oakes, L. M. (1992). Preschoolers' questions and parents' explanations: Causal thinking in everyday activity. *Cognitive Development, 7*(2), 213–233. https://doi.org/10.1016/0885-2014(92)90012-G.

Chin, C., & Osborne, J. (2008). Students' questions: A potential resource for teaching and learning science. *Studies in Science Education, 44*(1), 1–39. https://doi.org/10.1080/03057260701828101.

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis*. Hillsdale, NJ: Erlbaum.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334. https://doi.org/10.1007/bf02310555.

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academy Press.

Engel, S. (2011). Children's need to know: Curiosity in schools. *Harvard educational review, 81*(4), 625–645. https://doi.org/10.17763/haer.81.4.h054131316473115.

Engle, R. A., & Conant, F. R. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learners classroom. *Cognition and Instruction, 20*(4), 399–483.

Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research, 59*(2), 117–142. https://doi.org/10.3102/00346543059002117.

Forman, E. A., Engle, R. A., Venturini, P., & Ford, M. J. (2014). Introduction to special issue: International examinations and extensions of the productive disciplinary engagement framework. *International Journal of Educational Research, 64*, 149–155. https://doi.org/10.1016/j.ijer.2013.07.007.

Fornell, C. (1982). *A second generation of multivariate analysis methods*. New York: Praeger.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. https://doi.org/10.2307/3151312.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109.

Fredricks, J. A., Filsecker, M., & Lawson, M. A. (2016a). Student engagement, context, and adjustment: Addressing definitional, measurement, and methodological issues. *Learning and Instruction, 43*, 1–4. https://doi.org/10.1016/j.learninstruc.2016.02.002.

Fredricks, J. A., Wang, M.-T., Schall Linn, J., Hofkens, T. L., Sung, H., Parr, A., et al. (2016b). Using qualitative methods to develop a survey measure of math and science engagement. *Learning and Instruction, 43*, 5–15. https://doi.org/10.1016/j.learninstruc.2016.01.009.

Froiland, J. M., & Worrell, F. C. (2016). Intrinsic motivation, learning goals, engagement, and achievement in a diverse high school. *Psychology in the Schools, 53*(3), 321–336. https://doi.org/10.1002/pits.21901.

Greene, B. A., Miller, R. B., Crowson, H. M., Duke, B. L., & Akey, K. L. (2004). Predicting high school students' cognitive engagement and achievement: Contributions of classroom perceptions and motivation. *Contemporary Educational Psychology, 29*(4), 462–482. https://doi.org/10.1016/j.cedpsych.2004.01.006.

Gunuc, S., & Kuzu, A. (2015). Confirmation of campus-class-technology model in student engagement: A path analysis. *Computers in Human Behavior, 48*, 114–125. https://doi.org/10.1016/j.chb.2015.01.041.

Harrington, D. (2008). *Confirmatory factor analysis*. New York: Oxford University Press.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8.

Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education, 88*(1), 28–54. https://doi.org/10.1002/sce.10106.

Järvelä, S., Järvenoja, H., Malmberg, J., Isohätälä, J., & Sobocinski, M. (2016). How do types of interaction and phases of self-regulated learning set a stage for collaborative engagement? *Learning and Instruction, 43*, 39–51. https://doi.org/10.1016/j.learninstruc.2016.01.005.

Jeong, H., Songer, N. B., & Lee, S.-Y. (2007). Evidentiary competence: Sixth graders' understanding for gathering and interpreting evidence in scientific investigations. *Research in Science Education, 37*(1), 75–97. https://doi.org/10.1007/s11165-006-9014-9.

Jirout, J., & Klahr, D. (2012). Children's scientific curiosity: In search of an operational definition of an elusive concept. *Developmental Review, 32*(2), 125–160. https://doi.org/10.1016/j.dr.2012.04.002.

Justel, A., Peña, D., & Zamar, R. (1997). A multivariate Kolmogorov-Smirnov test of goodness of fit. *Statistics & Probability Letter, 35*(3), 251–259. https://doi.org/10.1016/S0167-7152(97)00020-5.

Kelloway, E. K. (1998). *Using LISREL for structural equation modeling: A researcher's guide*. Thousand Oaks, CA: Sage Publications.

Krajcik, J., Blumenfeld, P. C., Marx, R. W., Bass, K. M., Fredricks, J. A., & Soloway, E. (1998). Inquiry in project-based science classrooms: Initial attempts by middle school students. *Journal of the Learning Sciences, 7*(3&4), 313–350. https://doi.org/10.1080/10508406.1998.9672057.

Krajcik, J., & Czerniak, C. M. (2007). *Teaching children science in elementary and middle school classrooms: A project-based approach*. New York, NY: McGraw-Hill.

Krosnick, J. A., & Presser, S. (2010). Question and questionnaire design. In P. V. Marsden & J. D. Wright (Eds.), *Handbook of survey research* (2nd ed.). Bingley, UK: Emerald Group Publishing.

Kuh, G. D. (2003). What we're learning about student engagement from NSSE: Benchmarks for effective educational practices. *Education Collection, 35*(2), 24–32. https://doi.org/10.1080/00091380309604090.

Kuh, G. D. (2009). What student affairs professionals need to know about student engagement. *Education Collection, 50*(6), 683–706. https://doi.org/10.1353/csd.0.0099.

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. https://doi.org/10.1080/09500693.2015.1078521.

Li, Y. B., & Lerner, R. M. (2013). Interrelations of behavioral, emotional, and cognitive school engagement in high school students. *Journal of Youth and Adolescence, 42*(1), 20–32. https://doi.org/10.1007/s10964-012-9857-5.

Lin, C. C., & Tsai, C. C. (2008). Exploring the structural relationships between high school students' scientific epistemological views and their utilization of information commitments toward online science information. *International Journal of Science Education, 30*(15), 2001–2022. https://doi.org/10.1080/09500690701613733.

Linnenbrink-Garcia, L., Rogat, T. K., & Koskey, K. L. K. (2011). Affect and engagement during small group instruction. *Contemporary Educational Psychology, 36*(1), 13–24. https://doi.org/10.1016/j.cedpsych.2010.09.001.

Litman, J. A. (2005). Curiosity and the pleasures of learning: Wanting and liking new information. *Cognition and Emotion, 19*(6), 793–814. https://doi.org/10.1080/02699930541000101.

Luce, M. R., & Hsi, S. (2015). Science-relevant curiosity expression and interest in science: An exploratory study. *Science Education, 99*(1), 70–97. https://doi.org/10.1002/sce.21144.
Maxwell, S. E., & Delaney, H. D. (1993). Bivariate median splits and spurious statistical significance. *Psychological Bulletin, 113*(1), 181–190. https://doi.org/10.1037/0033-2909.113.1.181.
May, H. (2006). A multilevel Bayesian item response theory method for scaling socioeconomic status in international studies of education. *Journal of Educational and Behavioral Statistics, 31*(1), 63–79. https://doi.org/10.3102/10769986031001063.
McNeill, K. L., & Krajcik, J. (2008). Scientific explanations: Characterizing and evaluating the effects of teachers' instructional practices on student learning. *Journal of Research in Science Teaching, 45*(1), 53–78. https://doi.org/10.1002/tea.20201.
Mussel, P. (2010). Epistemic curiosity and related constructs: Lacking evidence of discriminant validity. *Personality and Individual Differences, 49*(5), 506–510. https://doi.org/10.1016/j.paid.2010.05.014.
National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.
National Research Council. (2000). *Inquiry and the national science education standards: A guide for teaching and learning*. Washington, DC: National Academy Press.
National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academies Press.
Nehring, A., Nowak, K. H., Zu Belzen, A. U., & Tiemann, R. (2015). Predicting students' skills in the context of scientific inquiry with cognitive, motivational, and sociodemographic variables. *International Journal of Science Education, 37*(9), 1343–1363. https://doi.org/10.1080/09500693.2015.1035358.
Newmann, F. M., Wehlage, G. G., & Lamborn, S. D. (1992). The significance and sources of student engagement. In F. M. Newmann (Ed.), *Student Engagement and Achievement in American Secondary Schools*. New York, NY: Teachers College Pre,ss.
Nystrand, M., & Gamoran, A. (1991). Instructional discourse, student engagement, and literature achievement. *Research in the Teaching of English, 25*(3), 261–290.
Pedaste, M., Mäeots, M., Siiman, L. A., Jong, T., van Riesen, S. A. N., Kamp, E. T., et al. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review, 14*, 47–61. https://doi.org/10.1016/j.edurev.2015.02.003.
Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher, 39*(2), 142–151. https://doi.org/10.3102/0013189x10363170.
Ryan, A. M., & Patrick, H. (2001). The classroom social environment and changes in adolescents' motivation and engagement during middle school. *American Educational Research Journal, 38*(2), 437–460.
Scheuermann, F., & Björnsson, J. (Eds.). (2009). *The transition to computer-based assessment: New approaches to skills assessment and implications for large-scale testing*. Luxembourg: European Communities.
Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research, 99*(6), 323–338. https://doi.org/10.3200/joer.99.6.323-338.
Sinha, T., Bai, Z., & Cassell, J. (2017). A new theoretical framework for curiosity for learning in social contexts. In É. Lavoué, H. Drachsler, K. Verbert, J. Broisin, & M. Pérez-Sanagustín (Eds.), *Proceedings of 12th European Conference on Technology Enhanced Learning* (pp. 254–269). Cham: Springer.
Spector, P. E., & Brannick, M. T. (2011). Methodological urban legends: The misuse of statistical control variables. *Organizational Research Methods, 14*(2), 287–305. https://doi.org/10.1177/1094428110369842.
Taiwan Ministry of Education. (2018). *Curriculum guidelines of 12-year basic education: Natural sciences*. Retrieved from https://www.naer.edu.tw/ezfiles/0/1000/attach/63/pta_18538_240851_60502.pdf
Wang, M.-T., Fredricks, J. A., Ye, F., Hofkens, T. L., & Linn, J. S. (2016). The math and science engagement scales: Scale development, validation, and psychometric properties. *Learning and Instruction, 43*, 16–26. https://doi.org/10.1016/j.learninstruc.2016.01.008.
West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling*. New York: Guilford Press.
Westland, J. C. (2010). Lower bounds on sample size in structural equation modeling. *Electronic Commerce Research and Applications, 9*(6), 476–487. https://doi.org/10.1016/j.elerap.2010.07.003.
Wu, M., & Adams, R. (2007). *Applying the Rasch model to psycho-social measurement: A practical approach*. Melbourne: Educational Measurement Solutions.

Wu, H.-K., & Huang, Y.-L. (2007). Ninth grade student engagement in teacher-centered and student-centered technology-enhanced learning environments. *Science Education*, *91*(5), 727–749. https://doi.org/10.1002/sce.20216.

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*(9–10), 1630–1650. https://doi.org/10.1080/09500693.2013.871660.

Wu, H.-K., Kuo, C. Y., Jen, T.-H., & Hsu, Y. S. (2015). What makes an item more difficult? Effects of modality and type of visual information in a computer-based assessment of scientific inquiry abilities. *Computers & Education*, *85*, 35–48. https://doi.org/10.1016/j.compedu.2015.01.007.

Wu, M., Tam, H.-P., & Jen, T.-H. (2016). *Educational measurement for applied researchers: Theory into practice*. Singapore: Springer Nature.

Wu, P. H., Kuo, C. Y., Wu, H.-K., Jen, T. H., & Hsu, Y. S. (2018). Learning benefits of secondary school students' inquiry-related curiosity: A cross-grade comparison of the relationships among learning experiences, curiosity, engagement, and inquiry abilities. *Science Education*, *102*(5), 917–950. https://doi.org/10.1002/sce.21456.

Zumbrunn, S., McKim, C., Buhs, E., & Hawley, L. R. (2014). Support, belonging, motivation, and engagement in the college classroom: A mixed method study. *Instructional Science, 42*(5), 661–684. https://doi.org/10.1007/s11251-014-9310-0.