

# THE EFFECTS OF A VIRTUAL LABORATORY AND META-COGNITIVE SCAFFOLDING ON STUDENTS' DATA MODELING COMPETENCES

Jeng-Fung Hung,  
Chun-Yen Tsai

## Introduction

The competences and knowledge of data modeling extensively guide contemporary humans in exploring the natural world and life. We live in a data-driven society, and people who read the newspaper carefully can see a variety of charts, tables, graphs, and other data representations (Doerr et al., 2017; English, 2012). With the availability of huge amounts of data, people can increasingly engage in democratic dialogue and public decision making, as well as participate in the discussion of science- and technology-related issues (Doerr et al., 2017; English, 2012; Provost & Fawcett, 2013). Data refers to things that can be recorded, analyzed, and reorganized, and presented in a quantitative format to organize the analysis so that people can use them to characterize the real world (Mayer-Schönberger & Cukier, 2014). Data modeling is a developmental process, starting with students' exploration and investigation of meaningful phenomena, identifying which attributes are worthy of attention in the phenomenon, and organizing, structuring, visualizing, and representing data for these attributes (English, 2012; Lehrer & Lesh, 2003; Lehrer & Schauble, 2000).

When students engage in scientific inquiry and communication in the science classroom, examining issues such as *what data can be used to answer this question, how to organize the data to explain the question clearly, or how to organize evidence to make the most convincing explanation* may provide opportunities for students to realize the process of data modeling (Namdar & Shen, 2015; NGSS Lead States, 2013). However, previous research (e.g., Srisawasdi & Panjaburee, 2019; Tsai, 2018) attached importance to inquiry teaching but failed to highlight the process and characteristics of data modeling. This may cause students to fail to develop data modeling competences (Jong et al., 2015; Nicolaou & Constantinou, 2014). For example, data collection usually requires the use of tools, but students may not be able to use such tools to measure data on things of interest (Michaels et al., 2007). In addition, when students interpret data, they may fail to grasp the meaning of the data as evidence of the hypothesis and may fail to interpret the meaning of the chart. Furthermore, students may impose rules on the data in a subjective way (Gott & Duggan, 2003). Data are not inherently structured. Scientists describe and organize data by selecting categories to impose structure on them. However,



JOURNAL  
OF BALTIC  
SCIENCE  
EDUCATION

ISSN 1648-3898 /Print/  
ISSN 2538-7138 /Online/

**Abstract.** *Previous studies on the effectiveness of virtual laboratories for learning have shown inconsistent results over the past decade. The purpose of this research was to explore the effects of a virtual laboratory and meta-cognitive scaffolding on students' data modeling competences. A quasi-experimental design was used. Three classes of eighth graders from southern Taiwan participated in this research and were assigned to the Experimental Group I (EG I), the Experimental Group II (EG II), and the Control Group (CG). EG I (n=25) received the virtual laboratory and meta-cognitive scaffolding in the teaching and learning. EG II (n=28) received the virtual laboratory only in the teaching and learning. The CG (n=27) received the lecture with the cookbook laboratory. The teaching unit was Heat and Specific Heat, and the teaching time for the three groups was six lessons (of 45 minutes each). The Data Modeling Competences Test (DMCT) designed by the research team was used as the data collection instrument. The results showed that the virtual laboratory and meta-cognitive scaffolding had effects on students' data modeling competences. This research shows the importance of the meta-cognitive scaffolding strategy for virtual laboratories when conducting data modeling teaching.*

**Keywords:** *data modeling, quasi-experimental design, meta-cognitive scaffolding, virtual laboratory*

Jeng-Fung Hung  
National Kaohsiung Normal University,  
Taiwan  
Chun-Yen Tsai  
National Sun Yat-sen University, Taiwan



students often fail to grasp this principle and tend to think that new problems can only be solved with new data, and rarely consider reviewing existing data to explore the potential problems (Michaels et al., 2007). Moreover, Lehrer and Schauble (2005) pointed out that students often have difficulty providing consistent data structures, and may not include important data, or may include redundant data during data analyses. Providing students with opportunities to organize and present data, in which they select, analyze and modify their data models, is an important way to solve these difficulties.

In several academic fields, laboratory work is required. In these fields, students are committed to practical problem solving and simulation training (Estriegana et al., 2019; Liu et al., 2017). In the laboratory learning environment, data and data analyses are needed, and play an increasingly important role. Data analyses include how to integrate, define, and transform data and how to process data. Data analysis can illustrate trends and important decision making about specific phenomena (Provost & Fawcett, 2013). If science teaching can help students focus more on how the data are generated and analyzed, they may better understand the data. Furthermore, students can use data to describe things, classify these things according to characteristics, seek their correlations, and finally explain the reasons for the phenomena. The learning growth resulting from these processes is an important goal of science education (Duschl et al., 2007), and learning to build scientific models and test them is also the focus of science education (Schwarz & White, 2005). In the model and modeling domain, Schwartz et al. (2009) advocated that modeling and the meta-modeling knowledge that make it meaningful must be put into practice together. In addition, related scholars (Brinson, 2015; Dori & Kaberman, 2012; Hodges et al., 2018; Wang et al., 2018) mentioned the benefits of virtual laboratory education. According to the theory of Vygotsky (1980), scaffolding is a timely auxiliary strategy that can help students achieve the potential development of higher level skills. This research aimed to explore the effects of a virtual laboratory and meta-cognitive scaffolding on students' data modeling competences.

## Literature Review

### *Data and Data Modeling*

Data are things that can be recorded, analyzed, and reorganized, while datafication of a certain phenomenon means presenting data in a quantitative format to facilitate analysis (Mayer-Schönberger & Cukier, 2014; Williamson, 2020). In order to obtain quantifiable information for datafication, it is necessary to know the measurement method and how to record the measured data. The abilities of measurement and recording promote the generation of datafication. Datafication enables people to use data to draw the true appearance of the world, grasp the key points in the huge amount of information, and obtain opportunities to solve problems (Mayer-Schönberger & Cukier, 2014). However, datafication is not just measurement and recording, but a process of data modeling (English & Sriraman, 2010). This process is based on determining what attributes affect the context and is related to answering questions (Guerrero-Ortiz et al., 2018). Students can collect data on these attributes and choose to characterize, organize, and display the data. They can then analyze the data and try to answer questions through reasoning.

The purpose of constructing data models is to characterize the data to make them more understandable (Guerrero-Ortiz et al., 2018; Shahbari & Peled, 2017). If students pay more attention to how the data are generated and the analyses of the data, they can better understand the data (Liu et al., 2017). First, students have to realize that data are used to answer questions. The question is what determines the type of information to be collected and the encoding and structure of the data (Michaels et al., 2007). Second, students need to understand that data are abstract because they represent observations of specific events (Liu et al., 2017). Interpretation of data can take many forms: a straight line distance can be expressed by several standard units, dynamic image recording can represent people's observations, or a reading on a thermometer can represent heat. In addition, students need to understand the concept of variables, including independent variables, dependent variables, and control variables (Liu et al., 2017), as well as the concepts of measurement, such as quantitative data, qualitative data, use of scale, and categorical and continuous variables (Organisation for Economic Co-operation and Development [OECD], 2013).

Data are not inherently structured and need to be interpreted. A structure must be imposed on the data. Scientists describe and organize data by selecting categories to impose structure on the data. In addition, data are represented in various ways to facilitate the understanding of different aspects of the studied phenomenon (Williamson, 2020). An important learning goal for students is to be able to understand the different types of conventions and attributes of data display (Liu et al., 2017). There are many different types of representations that can be used for data display, including tables, various types of graphics, and distribution. Interpreting data often



results in discovering and confirming correlations in the data (Shahbari & Peled, 2017), which may have different levels of complexity.

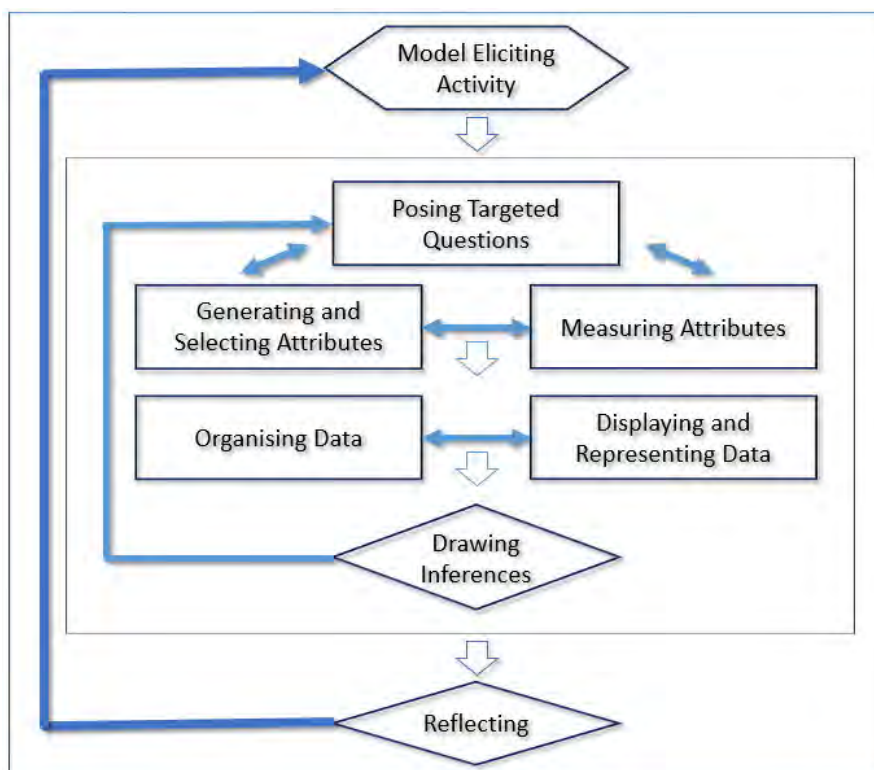
The data modeling competences that students may develop include (National Council of Teachers of Mathematics, 1989): 1) using measurements to describe and compare phenomena; 2) systematically collecting, organizing, and describing data; 3) constructing, reading, and interpreting tables, charts, and graphs; 4) analyzing tables and charts to determine attributes and correlations; 5) describing and characterizing correlations with tables, graphs and rules; 6) using tables, charts, text rules, and formulas to characterize the regularity of conditions and numbers, and to explore the interrelationship of these representations; and 7) analyzing functional correlations to explain how changes in one quantity cause changes in another. As the *big ideas* of science and mathematics refer to core concepts, principles, and theories of curricula (Lehrer & Schauble, 2000, 2005), data modeling should be an essential part of the science curriculum.

### A Teaching Model for Data Modeling

For students' inquiry learning, data modeling is a nested approach, and its inherent process can promote the development of students' ideas, models, and big ideas (Ärlebäck et al., 2015; Lehrer & Schauble, 2000). Ärlebäck et al. (2015) proposed an extended data-modeling approach (Figure 1) based on the data modeling model of English and Sriraman (2010). It is suggested that students be placed in a realistic and meaningful data modeling situation, where the starting point is the students' own problems. Based on the understanding of situations and problems, students discover solutions to problems by engaging in the inquiry process. According to contextualized questions, students must identify and decide which attributes affect the context. Students collect these attribute data and choose how to characterize, organize, and display the data. Finally, students analyze the data and answer their questions through reasoning (Ärlebäck et al., 2015; Lehrer & Schauble, 2000).

**Figure 1**

Process of the Extended Data-Modeling Approach (Ärlebäck et al., 2015)



This extended data-modeling approach begins with an initial model-eliciting activity which triggers students' ideas and models on a given learning goal (Årlebäck et al., 2015). The purpose of the model-eliciting activity is to place students in a meaningful context in which they face the need to develop strategies (Shahbari & Peled, 2017). Another purpose is to visualize students' previous experience and models (for themselves, peers, and teachers), and to clearly state the objects that can be reflected and discussed (Doerr & English, 2003; Lesh et al., 2003). In addition, the initial model-eliciting activity also helps students focus on asking questions when engaging in a model-exploration activity. In other words, the model-eliciting activity allows students to pool their thinking on the specific learning goals in a question which becomes more focused after discussion.

After the initial model-eliciting activities, model-exploration activities follow (Doerr & English, 2003). The main foci of these activities are to expose the underlying correlation structure of the elicited model. After students pose targeted questions in the model-exploration activity, they generate and measure attributes, organize and represent data, and draw inferences. The combination of model-eliciting and model-exploration activities lays the foundation for students to develop models, reasoning, and interpretation of the natural world (Årlebäck et al., 2015). Figure 1 also explicitly shows the reflecting stage of model-application activities. These activities are connected back to the students' original model-eliciting activity (Årlebäck et al., 2015). This activity provides students with an opportunity to reflect on their own methods and processes, as well as their applications in the real world, which is the key to modeling learning (Niss et al., 2007).

Årlebäck et al. (2015) pointed out that, from the perspective of models and modeling, learning is the development of models. Therefore, for students to learn how to model data, they need to develop their data model. The term *development* emphasizes the dynamic aspect of this process. In the process of the extended data-modeling approach, the students initiate and construct the model by participating in the model-eliciting activity. By carefully arranging model-exploration and model-application activities, students can use their models in other situations. Their models can be further developed, explored and applied. The extended data-modeling approach shown in Figure 1 provides theories and methods for the teaching and learning design of data modeling. Furthermore, from a learning viewpoint, the method of data modeling promotes the development of students' big concepts, and gives them first-hand experience (Lehrer & Schauble, 2000).

### *Virtual Laboratory and Science Learning*

The virtual laboratory is mainly based on computer software to simulate laboratory activities, so that learners can learn by doing experiments at any time and place (Estriegana et al., 2019; Wolski & Jagodzinski, 2019). The virtual laboratory can enhance the accessibility of the experimental environment and provide some efficient and practical tools (Estriegana et al., 2019) for students to learn by themselves. Such spontaneous learning becomes meaningful learning, since students become active learners in this kind of environment. In addition, the virtual laboratory can also make abstract concepts in science easier to understand (Dori & Kaberman, 2012; Wolski & Jagodzinski, 2019). Students in the physics department may encounter problems related to the development process of conceptual models, and the virtual laboratory can provide concrete models for understanding (Husnaini & Chen, 2019). Furthermore, learning science through experimentation allows students to develop different laboratory and process skills (Dori & Kaberman, 2012; Durand et al., 2019; Wolski & Jagodzinski, 2019). In a virtual laboratory, students can learn these basic skills (e.g., planning experiments, operating equipment, and recording data) and thinking first, so that they can use them later in an actual laboratory. When students re-experience these similar experimental environments, they can demonstrate better operational skills (Wolski & Jagodzinski, 2019).

Related scholars (Brinson, 2015; Hodges et al., 2018; Kolloffel & de Jong, 2013) have compared traditional laboratory (physical laboratory) and virtual laboratory learning, and have found that students' learning effectiveness was better in virtual laboratories than in traditional laboratories. However, Durand et al. (2019) and Quinn et al. (2009) compared the students' learning in biological laboratories in the two kinds of laboratories and found that students' learning outcome was better in traditional laboratories than in virtual laboratories. The research of Husnaini and Chen (2019) presented more in-depth findings. They found that students demonstrated better exploration skills in traditional laboratories while virtual laboratories were more effective than traditional laboratories in terms of understanding difficult concepts and improving students' self-efficacy (Husnaini & Chen, 2019).

From the above discussion, it can be seen that the effect of virtual laboratories on learning outcomes still needs further exploration. Some studies (Brinson, 2015; Hodges et al., 2018; Kolloffel & de Jong, 2013) revealed



that virtual laboratories were more effective, while others (e.g., Durand et al., 2019; Quinn et al., 2009) revealed that traditional laboratories were more effective. Students' learning in virtual laboratories may need to combine specific teaching strategies to strengthen the advantages of such laboratories. Some scholars (Nicolaou & Constantinou, 2014; Schwarz & White, 2005) believe that when modeling learning in virtual laboratories, meta-knowledge should be considered. Therefore, the following discussion will focus on the analysis of this teaching strategy.

### *Meta-knowledge and Meta-cognitive Scaffolding of Data Modeling*

Nicolaou and Constantinou (2014) reviewed previous studies and found that modeling competences could be divided into modeling practices and meta-knowledge. The modeling practices include the above-mentioned creating, revising, comparing, validating, and using models. The meta-knowledge includes the meta-modeling knowledge and the meta-cognitive knowledge of the modeling process. Meta-modeling knowledge refers to understanding the purpose of models and how to use and evaluate them (Nicolaou & Constantinou, 2014; Schwarz & White, 2005), while meta-cognitive knowledge refers to the competence of students to understand and reflect on the actual process of modeling (Abd-El-Khalick et al., 2004; Nicolaou & Constantinou, 2014; Schwarz & White, 2005). Meta-knowledge of modeling guides practice by helping students participate in modeling practice, and enables students to plan and evaluate their modeling practice (Schwarz & White, 2005).

The way to improve students' data modeling competences may provide the meta-cognitive scaffolding in practice. One way to scaffold students' modeling learning is to use scientific modeling criteria to promote students' reflection on the model (Cheng et al., 2017). Students' meaningful participation in modeling practice involves asking them to reflect the reasons and processes for managing modeling practice (Namdar & Shen, 2015). Schwartz et al. (2009) emphasized that modeling elements and meta-modeling knowledge are not independent learning goals, but it is the learning goals that should be integrated. The scientific modeling criteria can be obtained from the explanation of the concept and definition of evidence by Gott and Duggan (1995). They divided the concept of evidence into four parts: concepts related to *experimental design*, *measurement*, *data processing*, and *overall evaluation*. The first three items are directly related to data modeling, and the last one is related to the reliability and validity of the data modeling.

When modeling learning is in a digital environment, some of the advantages of computer software can be leveraged. The meta-cognitive scaffolding design proposed by Cheng et al. (2017) and White et al. (2009) can be referenced for the teaching of scientific inquiry: 1) Teachers can provide software advisors to define and model the inquiry processes, and suggest appropriate strategies to enable students to engage in inquiry; 2) Teachers can give students the opportunity to practice and control these processes when investigating in real inquiries; 3) Teachers can use scientific modeling criteria to enable students to monitor and reflect on their performance; and 4) By reflecting on their inquiry process, students can study and refine these processes. The suggestions of White et al. (2009) are applicable to the teaching and learning practice of data modeling.

### *Research Purpose*

Based on the above research background, this research aimed to explore the effects of a virtual laboratory and meta-cognitive scaffolding on students' data modeling competences. The data modeling teaching followed the processes of proposing the target question, generating and selecting variables, selecting measurement data, and organizing data models. The *Meta-cognitive Scaffolding Sheet* (see the *Instruments* section) with scientific modeling criteria suggested by Cheng et al. (2017) and White et al. (2009) was used as meta-cognitive scaffolding in the teaching process. The research purpose was to trigger, support, and promote students' competences in data modeling, which were divided into sub-tests (see the *Instruments* section). The following research questions were addressed:

- 1) Did the use of a virtual laboratory and meta-cognitive scaffolding cause differences in the total scores of the students' data modeling competences?
- 2) Did the use of a virtual laboratory and meta-cognitive scaffolding cause differences in the sub-test scores of the students' data modeling competences?
- 3) Did the use of a virtual laboratory and meta-cognitive scaffolding cause different performance in each competence of data modeling?



## Research Methodology

### *Research Design*

The quasi-experimental design was used in this research. Three eighth-grade classes were assigned as the Experimental Group I (EG I), the Experimental Group II (EG II), and the Control Group (CG). The students were arranged into classes in a random way when they enrolled in the school. EG I received the virtual laboratory and meta-cognitive scaffolding in the teaching process. EG II received the virtual laboratory in the teaching process. CG received the lecture with the cookbook laboratory in the teaching process (see the *Educational Design* section). The experiment period of the three groups was six lessons of 45 minutes each. The Test of Graphing Skills (see the *Instruments* section) was used as a pre-test before the experimental treatment. The Data Modeling Competences Test (see the *Instruments* section) was used as the post-test immediately after the experimental treatment.

### *Participants*

Three eighth-grade classes from a lower-secondary school in southern Taiwan were selected and assigned as EG I (n=25, 15 boys, 10 girls), EG II (n=28, 14 boys, 14 girls), and the CG (n=27, 15 boys, 12 girls). There were 80 students participating in this research, including 44 boys and 36 girls. The average age of these students was 13.5 years. Since the students were arranged into classes in a random way when they enrolled in the school, this research used the original class as the unit. EG I and EG II were taught by the same teacher Yeh, while the CG was taught by another teacher Zhang (both are pseudonyms). The two teachers had similar teaching experience (between 10 and 15 years of teaching science in lower-secondary school).

### *Teaching Context and the Interactive Data Modeling System*

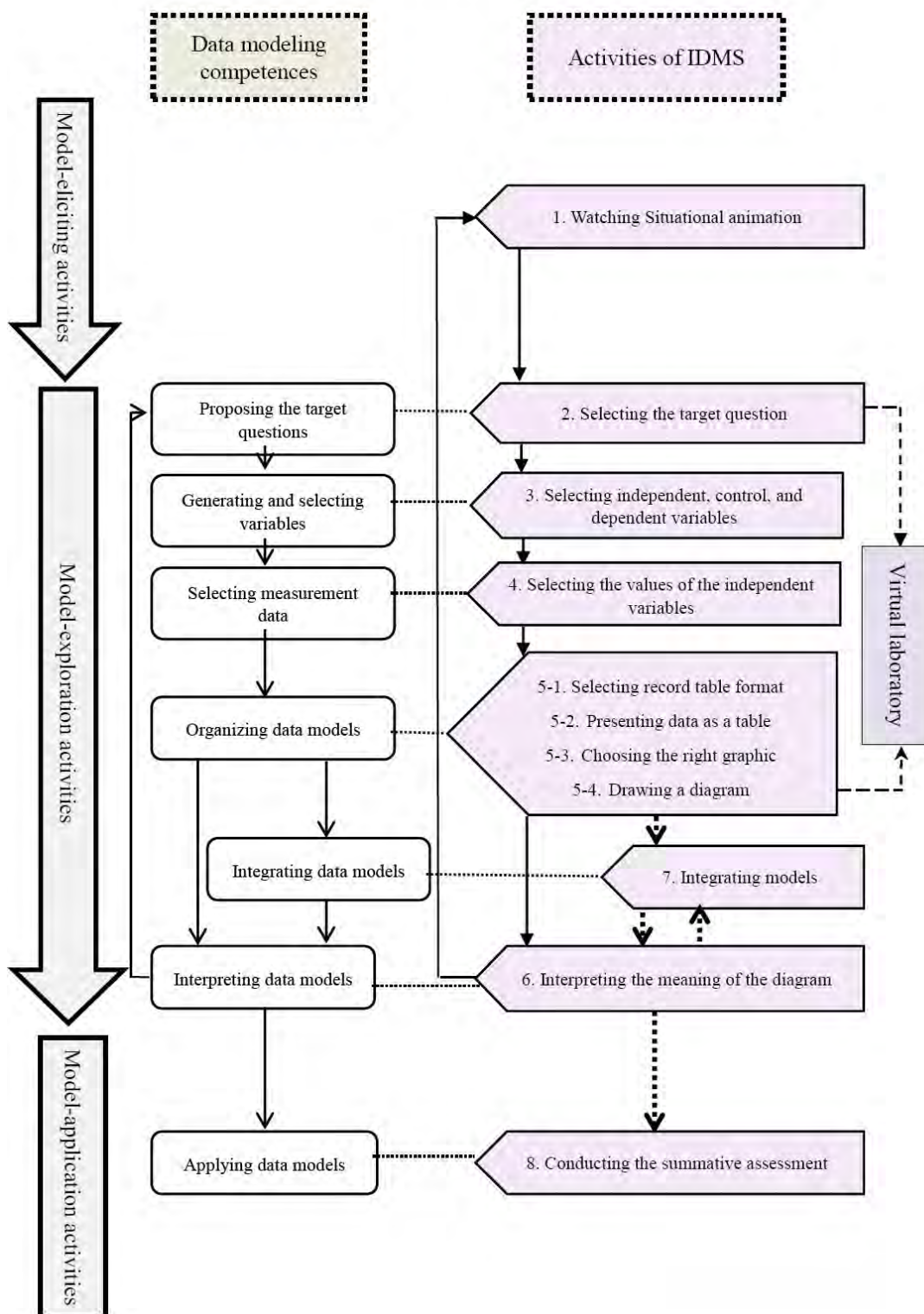
The purpose of the *Heat and Specific Heat* unit was to explore the temperature change of a substance after it is heated. Since this experiment involved the correlation between three variables, two sets of experiments had to be done. Then, the results of the two sets of experiments had to be integrated into a description of the correlation between the three variables:  $H = k \times m \times \Delta T$  (H is heat, k is the specific heat coefficient, m is the mass, and  $\Delta T$  is the temperature change). Although how the coefficient k is handled is the key, the textbook in Taiwan did not mention it and might mislead students in the process of learning how to model data. The course design of this research focused on the description of the correlation between *types of variables, experimental design, tables and graphs, and meaning of data*, as emphasized by Gott and Duggan (2003) and the OECD (2013), and the competences of data modeling were used as the basis for the teaching design.

Based on the extended data-modeling approach proposed by Ärlebäck et al. (2015), the data-modeling learning was divided into the three phases of model-eliciting, model-exploration, and model-application (Figure 2). In the model-eliciting activities, through the heat and specific heat-related situations that are common in daily life, the learning goal was to allow students to choose the questions to discuss. These questions involved changing the heating time and the mass of the water to measure the temperature change of the water. In the model-exploration activities, the learning objective was to allow the students to confirm the various variables based on the selected target question. Students measured the variables through virtual experiments, organized the data of the measurement results, and selected the correct graphics to display the data. Students then used this data model to infer and derive the correlations between the temperature change, heat change, and mass of water and glycerin. In the model-application activities, the learning goal was to further enable students to answer questions related to heat and specific heat.

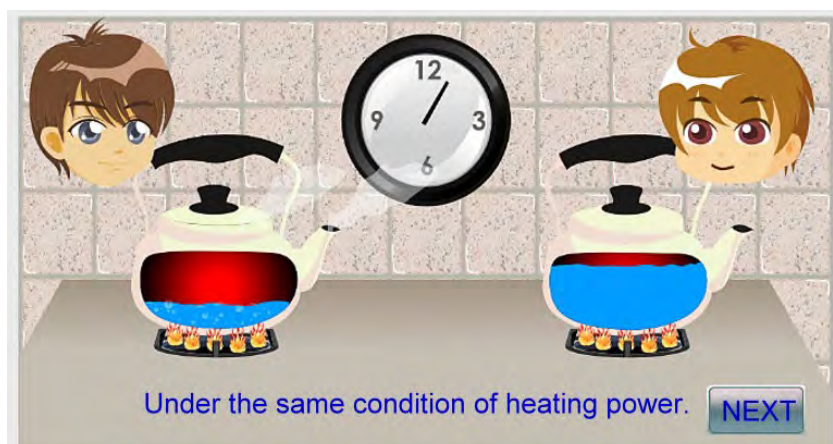
The activities of the *Interactive Data Modeling System* (IDMS) designed in this research mainly included the following steps: 1) watching situational animation; 2) selecting the target question; 3) selecting independent, control, and dependent variables; 4) selecting measurement data; and 5) organizing data models (Figure 2). In the *watching situational animation* step, situational animation presented life situations as trigger activities, and enabled students to focus on discovering issues that could be explored in the situation. This animation (Figure 3) shows two boys boiling water to make coffee, and the amount of water used by the two boys is different. Under the same condition of heating power provided by the gas stove, the smaller amount of water would boil first after heating for a period of time.



**Figure 2**  
Data Modeling Competences and Activities of IDMS













**Figure 3***The Situational Animation in the IDMS*

In the *selecting the target question* step, several questions that could be explored were provided for students to choose based on the phenomenon presented by the situational animation. Students had to confirm the following question to explore: *What is the correlation between the quality of water and the heating time of the water?* In the *selecting independent, control, and dependent variables* step, the students were asked to drag the variables of the experiment to the right positions according to the target question (Figure 4).

**Figure 4***The Selecting Independent, Control, and Dependent Variables Step in the IDMS*

Please make a correct experimental design for the target question.  
Please drag the variables on the left to the appropriate space on the right.

Target question: What is the correction between water quality and heating time?

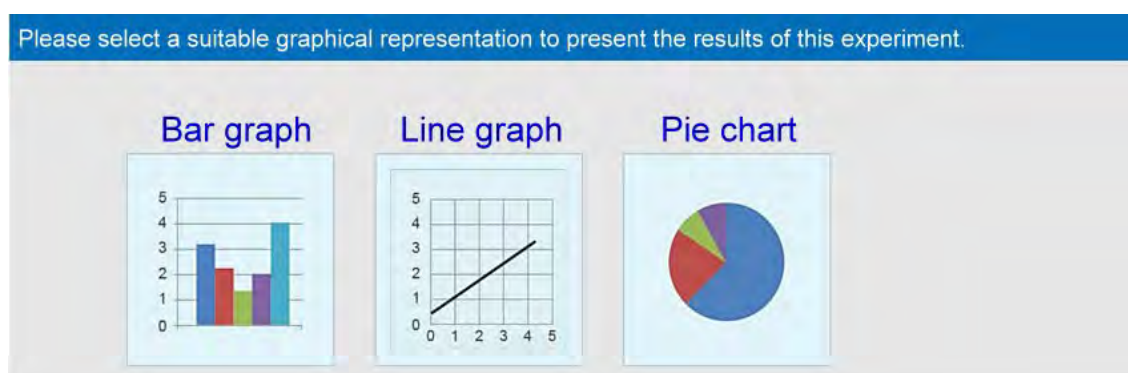
|   |  |                       |  |
|---|--|-----------------------|--|
| <br>Water quality (g)                | <br>Heating time (min)                      | Independent variables |  |
| <br>Initial temperature of water (C) | <br>End temperature of water (C)            | Control variables     |  |
| <br>Water temperature change (C)     | <br>Provided calories at a fixed time (cal) | Dependent variables   |  |
| <br>Mass of beaker (g)               | <br>Types of thermometers                   |                       |  |



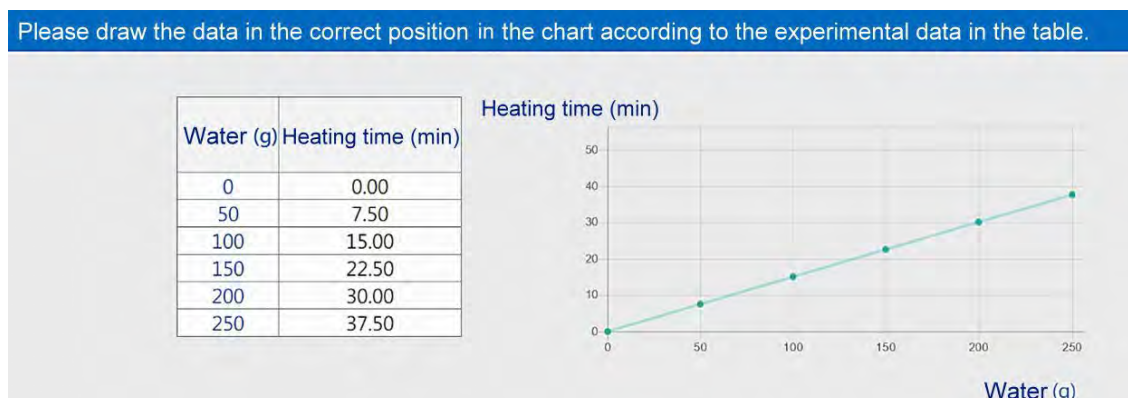
In the *selecting measurement data* step, the students had to manipulate the variables in the virtual experiment to obtain relevant data. In the *organizing data models* step, the students had to select a suitable graphical representation to model the data in the table according to the type of variable (Figure 5). They then had to place the dependent variable and independent variable on the correct position of the coordinate axis. They then drew a graph of the correlation between the two variables based on the data obtained from the virtual experiment (Figure 6). Through the process of data modeling, the students finally derived the correlation between the heat change, mass, specific heat, and temperature change of the substance.

**Figure 5**

*Selecting a Suitable Graphical Representation in the Organizing Data Models' Step*

**Figure 6**

*The Organizing Data Models in the IDMS Step*



### Educational Design

*EG I: the virtual laboratory and meta-cognitive scaffolding.* The teaching of EG I was as follows: (1) The teacher helped the students to operate IDMS, engage in inquiry, and answer questions raised in accordance with various steps; (2) The teacher recorded the students' responses on the whiteboard, and invited them to articulate the reasons for their responses. Students were asked to discuss and explain which model is more reasonable; (3) The teacher provided students with the *Meta-cognitive Scaffolding Sheet* (see the *Instruments* section) to enable them to monitor and reflect on their performance; (4) The teacher allowed the students to discuss in class to exchange ideas with each other to enable them to refine the process of data modeling.

*EG II: the virtual laboratory.* The teaching of EG II was as follows: (1) The teacher helped students to operate IDMS, engage in inquiry, and answer questions raised in accordance with various steps; (2) The teacher recorded the students' responses on the whiteboard, and invited them to articulate the reasons for their responses. Students



were asked to discuss and explain which model is more reasonable; (3) The teacher allowed the students to discuss in class to exchange ideas with each other to enable them to refine the process of data modeling. The difference between EG I and EG II teaching was that teaching for EG I provided students with the *Meta-cognitive Scaffolding Sheet*, while that for EG II did not.

*CG: lecture with cookbook laboratory.* The teaching for CG was carried out in the form of a lecture. The teacher prepared the slides for teaching. The slides were presented in the form of pictures and texts, during which there was an interactive questioning process between the teacher and students. In the experimental class, the students followed the experimental steps of the textbook in order. After finishing the experiment, they recorded the experiment data and results in the experiment record books, and answered the questions raised in the textbook. Then the students' record books were scored by the teacher. The teacher then discussed the problems in the record books with the students.

### Instruments

*Test of Graphing Skills.* The *Test of Graphing Skills* was designed by Su (2015) according to the competence index of the Test of Graphing in Science proposed by McKenzie and Padilla (1986), and was divided into two sub-tests of *Graphic Construction* (7 items) and *Graphic Interpretation* (8 items). There was a total of 15 items in the test, with one point for each correct answer, giving a total of 15 points. This test was reviewed by experts to establish content validity and tested by 63 lower-secondary school students with KR20 reliability of .85. In this research, the KR20 reliability of this test was .76.

*Data Modeling Competences Test (DMCT).* The DMCT was designed by referring to studies by English and Sriraman (2010), Årlebäck et al. (2015), and Gott and Duggan (2003), and the data modeling was divided into several competences. The DMCT is shown as Table 1 based on the competence index of McKenzie and Padilla's (1986) Test of Graphing in Science. After the DMCT was compiled, it was reviewed by two science educators and six science teachers in lower-secondary schools to establish content validity. The participants of the pilot study were 240 lower-secondary school students from two urban schools and a suburban school in southern Taiwan. The results showed that internal consistency reliability KR20 was .883 and the re-test reliability was .761. This test consisted of a total of 23 items in six item groups with a full score of 23 points. Each item group was in an independent context, and each item in the item group was an independent multiple-choice format.

**Table 1**  
*The DMCT Sub-scale and Competence Index*

| Sub-scale                             | Sub-competence                         | No. of items | Competence index   |
|---------------------------------------|--|--------------|--|
| Building Data Models                  | Proposing the target questions         | 4            | 1. A situation is given, and students can propose the target questions from the situational description.<br>2. Students can know the target questions of the experiment through the given data model.<br>3. Students can know the target questions of the experiment through the given experiment results and data.  |
|                                       | Generating and selecting variables     | 4            | 1. Students can find the appropriate independent, control, and dependent variables for the experiment from the questions they want to explore.<br>2. Students can identify the independent, control, and dependent variables when the data were obtained in the experiment.  |
|                                       | Selecting measurement data             | 4            | Students can choose the most appropriate quantity, range, and interval of units to collect data according to the problems they want to explore.  |
|                                       | Organizing and integrating data models | 4            | 1. Students can determine the types of variables that should be placed on the x and y axes.<br>2. Students can choose the correct data model according to the experimental data table.<br>3. Students can integrate the two related data models into a suitable data model.  |
| Interpreting and Applying Data Models | Interpreting and applying data models  | 7            | 1. Students can interpret the meaning of the data model.<br>2. Students can predict the data presented in the non-data model according to the trends of the data model and use the interpolation method.<br>3. Students can draw visual representation of the data obtained from the experiment on the given data model.<br>4. Students can perform simple calculations based on the data model. |



*Meta-cognitive Scaffolding Sheet.* According to Cheng et al.'s (2017) and White et al.'s (2009) suggestions for the meta-cognitive scaffolding proposed in the teaching of modeling and scientific inquiry, learning sheets were used to reflect scientific modeling criteria in this research. After each activity in Figure 2, the teacher asked the students to check the corresponding prompts of the learning sheets to promote their reflection (Table 2). For example: After the activity of *proposing the target questions*, the prompt that the students checked is, *Research questions are written in the form of questions that may have a correlation between the independent and dependent variables. The standard form: What is the effect of changing the 'acidity and alkalinity of water' (independent variable) on 'the growth of wheat seedlings' (dependent variable).*

**Table 2**  
*Reflective Prompts to Scaffold Students' Model Construction and Revision*

| Criteria for reflection on explanatory models | Reflective prompts   |
|---|--|
| Proposing the target questions                | Research questions are written in the form of questions that may have a correlation between the independent and dependent variables.   |
| Generating and selecting variables            | A controlled experiment is an experiment in which only one variable is changed at a time, and the other variables remain unchanged.  |
| Selecting measurement data                    | If the correlation between the two variables is a straight line, at least five different sets of experimental data are needed to make the graph. Too little data may misjudge the trend of the experimental results. |
| Organizing and integrating data models        | The independent variables are distributed along the horizontal axis, and the dependent variables are distributed upward along the longitudinal axis.   |

### *Data Analyses*

In this research, the Test of Graphic Skills was used as a covariant to perform the one-way analysis of covariance (ANCOVA). When the ANCOVA reached significance, the Scheffé's method was used for further comparison. In addition, the significance of each statistical test in this research was set to  $\alpha = .05$ . Moreover, in order to further explore the effect of the experimental treatment, the experimental effect index  $\eta^2$  proposed by Cohen (1988) for ANCOVA was adopted. The large, medium, and small effect sizes were .137, .058, and .010, respectively. Finally, in order to understand the learning problems of data modeling that existed in the three groups of students after the teaching process, the criteria of learning outcomes were based on the definition of Linn and Miller (2005): (1) If the average correct rate of the test item was less than 65%, it was considered as poor learning performance; (2) If the average correct rate of the test item was higher than 75%, it was considered as proficient learning performance; and (3) If the average correct rate of the test item was between 65% and 75%, it was considered as ordinary learning performance.

### **Research Results**

#### *The Difference in the Total Score of the DMCT*

The *Test of Graphic Skills* was used as a covariant, and the total score of DMCT was used as a dependent variable. The one-way covariate analysis was conducted to determine whether the three groups of students had differences in the total score of DMCT. The homogeneity test result of the regression coefficients (group x covariant) showed that the  $F$  value was 1.112 ( $p = .334 > .05$ ), which did not reach significance. This means that the test is in line with the assumption of homogeneity of the regression coefficients within the groups, and the ANCOVA can be performed.

The ANCOVA excluded the effect of the covariant on the dependent variable. The  $F$  value of the difference test was 3.988 and reached significance ( $p = .023 < .05$ ). The post-comparison results by Scheffé's method showed that the DMCT score of EG I was better than that of the CG ( $p = .006 < .05$ ); however, the DMCT score of EG I and that of EG II did not differ significantly, and there was no difference between the DMCT score of EG II and that of



the CG. In addition, the effect size of experimental treatment was  $\eta^2 = .10 > .058$ , and it had a medium effect size. Table 3 is the summary table of the mean score, standard deviation (SD), and adjusted mean score of the DMCT of three groups.

**Table 3**  
*The Performance of the DMCT of the Three Groups*

| Test | Group | <i>n</i> | <i>M</i> | <i>SD</i> | Adjusted mean |
|------|-------|----------|----------|-----------|---------------|
| DMCT | EG I  | 25       | 16.160   | 4.836     | 5.593         |
|      | EG II | 28       | 15.786   | 5.527     | 4.376         |
|      | CG    | 27       | 13.667   | 4.690     | 2.881         |

#### *The Difference in the Sub-tests of the DMCT*

There were two tests in this section. First, the sub-test *Graphic Construction* of the *Test of Graphing Skills* was used as a covariant and the sub-test *Building Data Models* of DMCT was used as the dependent variable. Next, the sub-test *Graphic Interpretation* of the *Test of Graphing Skills* was used as a covariant and the sub-test *Interpreting and Applying Data Model* of DMCT was used as the dependent variable. The one-way ANCOVA was conducted to determine whether the three groups of students had differences in the dependent variables. The homogeneity test results of the regression coefficients (group x covariant) showed that the *F* values were .527 ( $p = .593 > .05$ ) and .300 ( $p = .742 > .05$ ), which did not reach significance.

In the *Building Data Models* sub-test, the *F* value of the difference test of the groups on the dependent variable was 5.296 and reached significance ( $p = .007 < .05$ ). Post-hoc comparisons by Scheffé's method showed that EG I was better than the CG in the competence of *Building Data Models* ( $p = .002 < .05$ ); however, the *Building Data Models* score of EG I and that of EG II did not differ significantly, and there was no difference between the *Building Data Models* score of EG II and that of the CG. In addition, the effect size of experimental treatment was  $\eta^2 = .122 > .058$ , which was a medium to large effect. In the *Interpreting and Applying Data Models* sub-test, the *F* value of the difference test of the groups on the dependent variable was .536, and it did not reach significance ( $p = .587 > .05$ ). Table 4 is a summary table of the mean, SD, and adjusted mean of the two sub-tests of the three groups of students.

**Table 4**  
*The Performance of the Three Groups in the Two Sub-tests of DMCT*

| Sub-test                              | Group | <i>n</i> | Mean   | <i>SD</i> | Adjusted mean |
|---------------------------------------|-------|----------|--------|-----------|---------------|
| Building Data Models                  | EG I  | 25       | 11.560 | 3.124     | 7.088         |
|                                       | EG II | 28       | 11.179 | 3.945     | 5.977         |
|                                       | CG    | 27       | 9.630  | 3.364     | 4.508         |
| Interpreting and Applying Data Models | EG I  | 25       | 4.600  | 1.826     | 1.974         |
|                                       | EG II | 28       | 4.607  | 2.043     | 1.872         |
|                                       | CG    | 27       | 4.037  | 1.786     | 1.563         |

#### *The Problems Existing in Each Competence of the DMCT*

Table 5 is a summary table of the average correct rate and performance of each competence of data modeling. From the correct rate of the three groups in each competence, the possible problems of students in each competence of data modeling may be realized. Among the various competences of data modeling, the average correct rate of EG I and EG II in competence of *organizing and integrating data models* was greater than 75%, reach-



ing the proficient learning performance. The average answer rates of the students in EG I and EG II in the other four competences of data modeling were between 65% and 75%, only reaching ordinary learning performance. In the CG, the average correct rates of the other four competences of data modeling were all less than 65%. This result showed that the students in the CG had poor learning performance in most competences of data modeling.

**Table 5**

*The Performance of Data Modeling Competences of the Three Groups*

| Competence                             | Groups | Average correct rate (%) | Performance |
|--|--------|--------------------------|-------------|
| Proposing the target questions         | EG I   | 67                       | ++          |
|  | EG II  | 65                       | ++          |
|  | CG     | 61                       | +           |
| Generating and selecting variables     | EG I   | 68                       | ++          |
|  | EG II  | 66                       | ++          |
|  | CG     | 54                       | +           |
| Selecting measurement data             | EG I   | 73                       | ++          |
|  | EG II  | 72                       | ++          |
|  | CG     | 59                       | +           |
| Organizing and integrating data models | EG I   | 81                       | +++         |
|  | EG II  | 77                       | +++         |
|  | CG     | 67                       | ++          |
| Interpreting and applying data models  | EG I   | 66                       | ++          |
|  | EG II  | 66                       | ++          |
|  | CG     | 58                       | +           |

Note: + average correct rate < 65%; ++ 65% < average correct rate < 75%; +++ average correct rate > 75%

## Discussion

The results of this research reveal that there was no difference between the virtual laboratory (EG II) and the lecture with the cookbook laboratory (CG) in terms of improving the students' data modeling competences. The current research had similar findings to those of Husnaini and Chen (2019), who showed that the effects of virtual laboratories might not be evident in learning outcomes compared with those of physical laboratories. However, this research also found that the use of the virtual laboratory and meta-cognitive scaffolding in data modeling teaching (EG I) was more effective in terms of improving students' data modeling competences than the lecture with the cookbook laboratory (CG). This kind of research combining a virtual laboratory and the meta-cognitive scaffolding strategy has rarely been explored in previous studies. Such a finding is an insightful contribution provided by the current research to this field. This research was based on the learning scaffolding design proposed by White et al. (2009) for the teaching of meta-cognition in the context of scientific inquiry. The teaching effect with meta-cognitive scaffolding embedded in the virtual laboratory to improve the competences of students in data modeling was positive. If the teaching was not combined with meta-cognitive scaffolding, there was no evident difference in improving the competences of students' data modeling compared with the lecture with the cookbook laboratory. The above shows the importance of the meta-cognitive scaffolding strategy for virtual laboratory teaching.

Due to the complexity of the scientific inquiry process, students need to acquire meta-cognitive skills to facilitate their successful scientific inquiry learning (White et al., 2009). White et al. (2009) pointed out that students often do not have these meta-cognitive skills to effectively control or regulate the inquiry learning process. Therefore, learning scaffolding or teaching strategies that support metacognition is worth emphasizing. Reflective prompts to scaffold students' model construction and revision were provided in this study and might have been an effective learning aid. In this learning process, students' self-regulation might promote the management of



complex scientific inquiry learning (White et al., 2009). For example, in the competence of *generating and selecting variables*, the Meta-cognitive Scaffolding Sheet in this research suggested that students should only change one variable at a time as an independent variable. This allowed students to be aware of and control other variables during the process of data modeling.

Regarding the sub-tests, the comparison of the performance of the three groups of students in the competence of the *Building Data Models* after teaching was similar to the total score of DMCT. However, there was no difference between the three groups in the competence of *Interpreting and Applying Data Models*. This result shows that the data modeling teaching embedding a virtual laboratory and meta-cognitive scaffolding has a positive effect on training students to build data models. These competences included *proposing the target questions, generating and selecting variables, selecting measurement data, and organizing and integrating data models*. Such results are consistent with part of Schwarz and White's (2005) research, which revealed that students could develop several data modeling competences after using meta-cognitive strategies in modeling teaching. Schwarz and White (2005) stated that the use of meta-cognitive strategies in modeling teaching could enable students to develop epistemologies of science and understand that models in the experimental process mainly simulate and predict real-world phenomena. Such learning may allow students to be more aware of and in control of the learning process and handle the learning of modeling competences. However, the competence of *interpreting and applying data models* belongs to the more advanced data modeling competences, and may not be effectively improved by relying solely on the data modeling teaching with the cognitive scaffolding. If this competence is to be developed, other teaching strategies may be combined to strengthen students' competence of *interpreting and applying data models*.

From the descriptive statistical analysis of various data modeling competences, the average answer rate of the students in the two experimental groups (EG I and EG II) in the competence of *organizing and integrating data models* was greater than 75%, which can be considered as proficient learning performance. The students in the two experimental groups exhibited ordinary learning performance in the other four data modeling competences. The current research had similar findings to those of Dori and Kaberman (2012), who showed the effects that virtual laboratories might have on students' modeling sub-skills. The reason may be that the competence of *organizing and integrating data models* focused on deriving  $H = m \times s \times \Delta T$  from data and charts, and this was the learning focus of the *Heat and Specific Heat* unit. The learning of this competence might have attracted students' attention more than the other four competences. As for the CG, only the competence of *organizing and integrating data models* was considered as ordinary learning performance, while the other four competences were considered as poor learning performance. This shows the shortcomings of lectures with cookbook laboratory teaching for enhancing students' data modeling competences. Since this inference was made from the analysis of descriptive statistics, it cannot be generalized to a wider population, and may be confirmed by further research.

One difficulty that students may encounter in learning science is learning abstract concepts. Dori and Kaberman (2012) found that students became more proficient in writing molecular structural formulas and drawing models in space after virtual laboratory teaching. They also found that students could establish connections between different representations of molecules and improve their understanding of chemistry. This means that students can learn the skills of symbol transformation and the transformation between micro and macro concepts in the virtual laboratory environment. In other words, the virtual laboratory learning environment may promote the development of students' modeling competences to transform abstract representations and micro concepts into concrete and macro concepts.

In the cookbook laboratory, students can only do experiments according to the textbook. During the experiment, the equipment in the lower-secondary school classroom may not be very accurate and may cause some experimental errors (Husnaini & Chen, 2019). Students may also spend a lot of time making experimental records. The interference of these experimental operations may affect their data modeling learning. From the perspective of cognitive loading, physical laboratories may only promote students' basic inquiry ability (Husnaini & Chen, 2019). On the other hand, students can concentrate on repeating the experimental process in the virtual laboratory, and observe the changes of various variables. This simpler simulation environment may allow them to focus on the establishment of data models, and then develop various data modeling competences.

## Conclusions

This research adopted the extended data-modeling approach and meta-cognitive scaffolding design to construct IDMS, and conducted data modeling teaching. In terms of the performance of data modeling compe-





tences, the students using IDMS with the meta-cognitive scaffolding had better learning outcomes compared with the students who learned through the lecture and cookbook laboratory teaching. Without the meta-cognitive scaffolding, the data modeling teaching using the virtual laboratory had no difference from the lecture with the cookbook experimental teaching in terms of improving the performance of the students' data modeling competences. When the average correct rates of data modeling competences were carefully compared, it revealed that the two experimental groups achieved ordinary learning performance in four data modeling competences, while the control group had poor learning performance in four competences.

This research revealed that the virtual laboratory had learning effects on students' data modeling competences. In the future teaching design, it is suggested that the virtual laboratory may be conducted in the science classroom to enhance students' data modeling competences. In particular, this research revealed that the virtual laboratory with the meta-cognitive scaffolding might better show the effects on students' data modeling competences. In the future, it is suggested that the meta-cognitive scaffolding design may be incorporated into the virtual laboratory teaching to improve the teaching effects. Future research can use different meta-cognitive scaffoldings in the virtual laboratory to confirm the effects of different meta-cognitive scaffoldings on students' learning outcomes.

There are two limitations of this research that should be noted. First, the number of participants in each group was less than 30. This might make it less likely that the statistical test of the differences between the groups reached a significant level. Second, the data modeling learning of the virtual laboratory in this research focused on the *Heat and Specific Heat* unit; generalization to other scientific concepts may need to be further verified.

## References

- Abd-El-Khalick, F., Boujaoude, S., Duschl, R., Lederman, N. G., Mamlok-Naamn, R., Hofstein, A., Niaz, M., Treagust, D., & Tuan, H. L. (2004). Inquiry in science education: International perspectives. *Science Education*, 88(3), 397-419. <https://doi.org/10.1002/sce.10118>
- Årlebäck, J. B., Blomberg, P., & Nilsson, P. (2015). An instructional design perspective on data modelling for learning statistics and modelling. In O. Helenius et al. (Eds.), *Development of mathematics teaching: Design, scale, effects* (pp. 37-46). Swedish Society for Research in Mathematics Education. <http://liu.diva-portal.org/smash/record.jsf?pid=diva2%3A917833&dsid=9974>
- Brinson, J. R. (2015). Learning outcome achievement in non-traditional (virtual and remote) versus traditional (hands-on) laboratories: A review of the empirical research. *Computers & Education*, 87, 218-237. <https://doi.org/10.1016/j.compedu.2015.07.003>
- Cheng, M. F., Lin, J. L., Lin, S. Y., & Cheng, C. H. (2017). Scaffolding middle school and high school students' modeling processes. *Journal of Baltic Science Education*, 16(2), 207-217. <http://www.scientiasocialis.lt/jbse/?q=node/559>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Ed.). Routledge.
- Doerr, H. M., Årlebäck, J. B., & Misfeldt, M. (2017). Representations of modelling in mathematics education. In: G. Stillman, W. Blum, & G. Kaiser (Eds.), *Mathematical modelling and applications. International perspectives on the teaching and learning of mathematical modelling* (pp. 71-81). Springer. [https://doi.org/10.1007/978-3-319-62968-1\\_6](https://doi.org/10.1007/978-3-319-62968-1_6)
- Doerr, H. M., & English, L. D. (2003). A modeling perspective on students' mathematical reasoning about data. *Journal for Research in Mathematics Education*, 34(2), 110-136. <https://doi.org/10.2307/30034902>
- Dori, Y., & Kaberman, Z. (2012). Assessing high school chemistry students' modeling sub-skills in a computerized molecular modeling learning environment. *Instructional Science*, 40(1), 69-91. <https://doi.org/10.1007/s11251-011-9172-7>
- Durand, M. T., Restini, C. B. A., Wolff, A. C. D., Faria, M. Jr., Couto, L. B., & Bestetti, R. B. (2019). Students' perception of animal or virtual laboratory in physiology practical classes in PBL medical hybrid curriculum. *Advances in Physiology Education*, 43, 451-457. <https://doi.org/10.1152/advan.00005.2019>
- Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (2007). *Taking science to school: Learning and teaching science in grades K-8*. The National Academies Press. <https://eric.ed.gov/?id=ED536343>
- English, L. D. (2012). Data modelling with first-grade students. *Educational Studies in Mathematics*, 81, 15-30. <https://doi.org/10.1007/s10649-011-9377-3>
- English, L., & Sriraman, B. (2010). Problem Solving for the 21<sup>st</sup> Century. In: B. Sriraman, & L. English (Eds), *Theories of mathematics education. Advances in mathematics education*. Springer. [https://doi.org/10.1007/978-3-642-00742-2\\_27](https://doi.org/10.1007/978-3-642-00742-2_27)
- Estrigana, R., Medina-Merodio, J., & Barchino, R. (2019). Student acceptance of virtual laboratory and practical work: An extension of the technology acceptance model. *Computers & Education*, 135, 1-14. <https://doi.org/10.1016/j.compedu.2019.02.010>
- Gott, R., & Duggan, S. (1995). *Investigative work in the science curriculum: Developing science and technology education*. Open University Press. <https://eric.ed.gov/?id=ED390642>
- Gott, R., & Duggan, S. (2003). *Understanding and using scientific evidence: How to critically evaluate data*. SAGE Publications. <https://eric.ed.gov/?id=ED482948>
- Guerrero-Ortiz, C., Mena-Lorca, J., & Soto, A. M. (2018). Fostering transit between real world and mathematical world: Some phases on the modelling cycle. *International Journal of Science and Mathematics Education*, 16, 1605-1628. <https://doi.org/10.1007/s10763-017-9856-9>



- Hodges, G. W., Wang, L., Lee, J., Cohen, A., & Jang, Y. (2018). An exploratory study of blending the virtual world and the laboratory experience in secondary chemistry classrooms. *Computers & Education*, 122, 179–193. <https://doi.org/10.1016/j.compedu.2018.03.003>
- Husnaini, S. J., & Chen, S. (2019). Effects of guided inquiry virtual and physical laboratories on conceptual understanding, inquiry performance, scientific inquiry self-efficacy, and enjoyment. *Physical Review Physics Education Research*, 15, 010119-1-16. <https://doi.org/10.1103/PhysRevPhysEduRes.15.010119>
- Jong, J.-P., Chiu, M.-H., & Chung, S.-L. (2015). The use of modeling-based text to improve students' modeling competencies. *Science Education*, 99(5), 986-1018. <https://doi.org/10.1002/sce.21164>
- Kolloffel, B., & de Jong, T. (2013). Conceptual understanding of electrical circuits in secondary vocational engineering education: Combining traditional instruction with inquiry learning in a virtual lab. *Journal of Engineering Education*, 102(3), 375–393. <https://doi.org/10.1002/jee.20022>
- Lehrer, R., & Lesh, R. (2003). Mathematical learning. In W. M. Reynolds, & G. E. Miller (Eds.), *Handbook of psychology: Educational psychology* (Vol. 7, pp. 357–391). John Wiley. <https://psycnet.apa.org/record/2003-04684-015>
- Lehrer, R., & Schauble, L. (2000). Modeling in mathematics and science. In R. Glaser (Ed.), *Advances in instructional psychology* (Vo. 5, pp. 101–159). Lawrence Erlbaum Associates Publishers. <https://books.google.com/books?id=FGMmiG8cDxoC>
- Lehrer, R., & Schauble, L. (2005). Developing modeling and argument in the elementary grades. In T. A. Romberg, T. P. Carpenter, & F. Dremock (Eds.), *Understanding mathematics and science matters. Studies in mathematical thinking and learning series* (pp. 29–53). Lawrence Erlbaum Associates Publishers. <https://psycnet.apa.org/record/2005-02699-002>
- Lesh, R., Cramer, K., Doerr, H. M., Post, T., & Zawojewski, J. S. (2003). Model development sequences. In R. Lesh, & H. M. Doerr (Eds.), *Beyond constructivism: Models and modeling perspectives on mathematics problem solving, learning, and teaching* (pp. 35–58). Psychology Press. [https://books.google.com/books?id=0oX7\\_lkdyJMC](https://books.google.com/books?id=0oX7_lkdyJMC)
- Linn, R. L., & Miller, M. D. (2005). *Measurement and assessment in teaching*. Pearson. <https://books.google.com.tw/books?id=bE2dAAAAAAAJ>
- Liu, C. Y., Wu, C. J., Wong, W. K., Lien, Y. W., & Chao, T. K. (2017). Scientific modeling with mobile devices in high school physics labs. *Computers & Education*, 105, 44-56. <https://doi.org/10.1016/j.compedu.2016.11.004>
- Mayer-Schönberger, V., & Cukier, K. (2014). *Big data: A revolution that will transform how we live, work, and think*. John Murray. <https://books.google.com/books?id=uy4lh-WEhhIC>
- McKenzie, D. L., & Padilla, M. J. (1986). The construction and validation of the test of graphing in science (TOGS). *Journal of Research in Science Teaching*, 23(7), 571-579. <https://doi.org/10.1002/tea.3660230702>
- Michaels, S., Shouse, A. W., & Schweingruber, H. A. (2007). *Ready, set, science! Putting research to work in K-8 science classrooms*. National Academies Press. <https://books.google.com/books?id=ME-W-WRjKEgC>
- Namdar, B., & Shen, J. (2015). Modeling-oriented assessment in K-12 science education: A synthesis of research from 1980 to 2013 and new directions. *International Journal of Science Education*, 37(7), 993-1023. <https://doi.org/10.1080/09500693.2015.1012185>
- National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Author. <https://books.google.com.tw/books?id=UdHaAAAAAAAJ>
- NGSS Lead States. (2013). *Next generation science standards: For states, by states*. The National Academies Press. <https://eric.ed.gov/?id=ED555235>
- Nicolaou, C. T., & Constantinou, C. P. (2014). Assessment of the modeling competence: A systematic review and synthesis of empirical research. *Educational Research Review*, 13, 52-73. <https://doi.org/10.1016/j.edurev.2014.10.001>
- Niss, M., Blum, W., & Galbraith, P. L. (2007). Introduction. In W. Blum, P. L. Galbraith, H. -W. Henn, & M. Niss (Eds.), *Modelling and applications in mathematics education: The 14th ICMI study* (pp. 3–32). Springer. <https://doi.org/10.1007/978-0-387-29822-1>
- Organisation for Economic Co-operation and Development. (2013). *PISA 2015 Draft science framework*. Author. <https://www.oecd.org/pisa/pisaproducts/pisa2015draftframeworks.htm>
- Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*, 1(1), 51-59. <https://doi.org/10.1089/big.2013.1508>
- Quinn, J. G., King, K., Roberts, D., Carey, L., & Mousley, A. (2009). Computer based learning packages have a role, but care needs to be given as to when they are delivered. *Bioscience Education*, 14(1), 1-11. <https://doi.org/10.3108/beej.14.5>
- Schwartz, D., Chase, C., Chin, D. B., Oppezzo, M., Kwong, H., Okita, S., Biswas, G., Roscoe, R., Jeong, H., & Wagster, J. (2009). Interactive metacognition: Monitoring and regulating a teachable agent. In D. J. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Handbook of metacognition in education* (pp. 340-358). Routledge. <https://doi.org/10.4324/9780203876428>
- Schwarz, C. & White, B. (2005). Meta-modeling knowledge: Developing students' understanding of scientific modeling. *Cognition and Instruction*, 23(2), 165-205. [https://doi.org/10.1207/s1532690xci2302\\_1](https://doi.org/10.1207/s1532690xci2302_1)
- Shahbari, J. A., & Peled, I. (2017). Modelling in primary school: Constructing conceptual models and making sense of fractions. *International Journal of Science and Mathematics Education*, 15, 371–391. <https://doi.org/10.1007/s10763-015-9702-x>
- Srisawasdi, N., & Panjaburee, P. (2019). Implementation of game-transformed inquiry-based learning to promote the understanding of and motivation to learn chemistry. *Journal of Science Education and Technology*, 28(2), 152–164. <https://doi.org/10.1007/s10956-018-9754-0>
- Su, Y. R. (2005). *Investigating the relationships between instructional modules, graphing skills, and conceptual understandings: Eighth graders' learning about density*. Unpublished master's thesis, National Taiwan Normal University, Taiwan.
- Tsai, F. H. (2018). The development and evaluation of a computer-simulated science inquiry environment using gamified elements. *Journal of Educational Computing Research*, 56(1), 3-22. <https://doi.org/10.1177/0735633117705646>



- Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Harvard University Press. <https://books.google.com.tw/books?id=lrq913IEZ1QC>
- White, B., Frederiksen, J., & Collins, A. (2009). The interplay of scientific inquiry and metacognition: More than a marriage of convenience. In D. Hacker, J. Dunlosky, & A. Graesser, (Eds.), *Handbook of metacognition in education* (pp.175-205). Routledge. <https://doi.org/10.4324/9780203876428>
- Williamson, B. (2020). Datafication of education: A critical approach to emerging analytics technologies and practices. In H. Beetham, & R. Sharpe (Eds.), *Rethinking pedagogy for a digital age: Principles and practices of design* (pp. 212-226). Routledge. <https://doi.org/10.4324/9781351252805>
- Wolski, R., & Jagodzinski, P. (2019). Virtual laboratory—Using a hand movement recognition system to improve the quality of chemical education. *British Journal of Educational Technology*, 50(1), 218–231. <https://doi.org/10.1111/bjet.12563>

Received: July 06, 2020

Accepted: November 21, 2020

Cite as: Hung, J.-F., & Tsai, C.-Y. (2020). The effects of a virtual laboratory and meta-cognitive scaffolding on students' data modeling competences. *Journal of Baltic Science Education*, 19(6), 923-939. <https://doi.org/10.33225/jbse/20.19.923>

---

**Jeng-Fung Hung**

Professor, Graduate Institute of Science Education & Environmental Education, National Kaohsiung Normal University, No.62, Shenhong Rd., Yanchao Dist., Kaohsiung City 82446, Taiwan.  
Email: t1873@nknuc.nknu.edu.tw

**Chun-Yen Tsai**  
(Corresponding author)

Professor, Center for General Education, National Sun Yat-sen University, No.70, Lienhai Rd., Gushan Dist., Kaohsiung City 80424, Taiwan.  
E-mail: ctsai@mail.nsysu.edu.tw  
<https://orcid.org/0000-0002-4016-5614>

---

