



# Modelling Competence in Teacher Education: Comparing Meta-modelling Knowledge, Modelling Practices and Modelling Products Between Pre-service and In-service Teachers

Song Xue<sup>1</sup> · Keith Topping<sup>2</sup> · Elizabeth Lakin<sup>2</sup> · Moritz Krell<sup>3</sup>

Accepted: 20 June 2024

© The Author(s), under exclusive licence to Springer Nature B.V. 2024

## Abstract

There has been increased attention recently on models and modelling within the global science education field. Research has begun to skew towards a competence-based perspective of models and modelling, as teachers are experiencing challenges and do not have the required competence in modelling from either theoretical or practical perspectives. This study was designed to comparatively investigate pre-service science teachers' (PSTs) and in-service science teachers' (ISTs) modelling competence. A rating scale questionnaire was developed to assess meta-modelling knowledge. Additionally, a Black Box modelling task was designed to evaluate modelling practices and products by using two techniques: think-aloud and drawings. The resulting data was then coded and scored with validated rubrics. Quantitative analysis revealed that ISTs outperformed the PSTs in meta-modelling knowledge but they had an almost equal level in modelling practices and products, which were not at a satisfactory level. Furthermore, modelling practices and products were positively related, but no significant relationships were found between meta-modelling knowledge, modelling practices and products. Results of qualitative analyses further indicated higher-level practices were reflected in the analysis of correct model products, which was accompanied by sophisticated scientific knowledge and other advanced scientific skills. Implications of this study for science education research and teacher professional development are discussed.

**Keywords** Modelling practice · Modelling competence · Meta-modelling knowledge · Modelling product · Science teacher education

## Introduction

Over the past three decades, there has been substantial growth in the amount of research on models and modelling in science education (Chiu & Lin, 2022). Some international organizations and institutions have outlined the importance of scientific models and modelling practices in learning and teaching science (Ministry of Education, 2022; National Research Council, 2012; Nielsen & Nielsen, 2021). A new direction of research in this field started to focus on developing a competence-oriented perspective on modelling, specifically,

---

Extended author information available on the last page of the article

“modelling competence” (Upmeier zu Belzen et al., 2019). Due to the limited research on this focus, there is a continuing need for an operational framework of modelling competence to identify the components that could guide teaching and assessment, specifically process-oriented assessment (Göhner et al., 2022; Nicolaou & Constantinou, 2014; Nielsen & Nielsen, 2021). In addition, while the issue of how to assess modellers’ understanding of models and modelling practices has been discussed for three decades, the development of more assessments is still required and these must be validated for researchers to understand modelling competence better (Chiu & Lin, 2019; Nicolaou & Constantinou, 2014).

Because of the complexity of supporting students’ engagement in modelling, attention is being paid to understanding teachers’ modelling competence. Knowing about aspects of the modelling competence of both pre-service teachers (PSTs) and in-service teachers (ISTs) could lead to the further development of teacher education programmes (Göhner & Krell, 2022; Göhner et al., 2022). Much prior research focuses on teachers’ understanding of models and modelling (Danusso et al., 2010; Justi & Gilbert, 2002a, 2002b), rather than teachers’ capacities to engage in modelling practices and develop high-quality modelling products (Chiu & Lin, 2019; Göhner et al., 2022). This intricacy is deepened by the ongoing controversy over the relationship between the components of modelling competence, namely meta-modelling knowledge, modelling practices, and modelling products (Cheng & Lin, 2015; Göhner et al., 2022; Schwarz & White, 2005; Sins et al., 2009). However, a profound understanding of how these components interact suggests that a teacher’s meta-modelling knowledge significantly influences their classroom modelling practices, which then determine the quality of the resultant modelling products (Vo et al., 2015). The discrepancies between the broad theoretical frameworks and their classroom applications underscore the need for potential refinements in teacher education. Highlighting this, the need for empirical evidence becomes even more pronounced, emphasizing the necessity to gauge the real-world efficacy of modelling competencies. Given the evolving nature of modelling techniques and methodologies, it’s clear that the realm of teacher education is not a static field but an ongoing journey, advocating for the continuous professional development of both PSTs and ISTs.

The present study aims to outline a framework for conceptualizing the theoretical position of the three modelling competence components: meta-modelling knowledge, modelling practice, and the modelling product. Then the framework can be used as a theoretical foundation for the development and implementation of assessments for teachers’ modelling competence. Finally, the empirical analysis could provide valuable evidence for evaluating the framework and making the theoretical position operationalised and condensed.

In this study, modelling competence can be divided into three components: meta-modelling knowledge, modelling practice, and the modelling product. Two research questions will be addressed in this study:

1. What is the relationship between the three components of modelling competence?
2. What are the similarities and differences between PSTs and ISTs’ modelling competence components?

## Framework of modelling competence in science education

The term competence is described as “domain-specific cognitive dispositions that are required to successfully cope with certain situations or tasks, and that are acquired by learning processes” (Koeppen et al., 2008, p. 62). With a focus on models and modelling

in science education, modelling competence is adapted to Weinert's (2001) definition of competence, referring to successful mastery and reflecting sufficient knowledge and practice (for a particular aspect). Consequently, modelling competence reflects a person's potential by combining knowledge about models and modelling with modelling practice to produce a model that meets cognitive needs in specific content areas.

Different frameworks of modelling competence (FMC) were developed for science education purposes, to explore, "how models are used, why they are used, and what their strengths and limitations are, in order to appreciate how science works and the dynamic nature of knowledge that science produces" (Schwarz et al., 2009, pp. 634–635). The components of modelling competence are controversial. Gilbert and Justi (2016) proposed the approach emphasizes the use of models and modelling to enhance the understanding and competence of learners. By allowing students to develop their own models, it encourages active participation and establishes a creative learning environment. Nicolaou and Constantinou (2014) defined modelling competence into two broad categories, namely modelling practices, and meta-knowledge. This FMC attempts to support the claim that students' modelling competence can emerge as a result of active engagement in specific modelling practices and is shaped by meta-knowledge about models and modelling. Some studies promoted other components, modelling product (Chiu & Lin, 2019; Göhner et al., 2022; Namdar & Shen, 2015) and subject-specific knowledge (Nielsen & Nielsen, 2021). These various frameworks for how learners understand models and modelling and rationale for integrating models and modelling into their teaching have been discussed and refined.

Based on the existing FMCs, this presented study outlines a framework describing the theoretical position of three components: meta-modelling knowledge, modelling practice, and modelling product. As shown in Fig. 1, modelling products can be erroneous, partially correct, or correct. Meta-modelling knowledge includes four dimensions: knowledge about the modelling process, the nature of models, types of models, and the purpose of models. There are twelve different modelling practices distributed across different stages of the modelling process (e.g., comparing the experimental data, generating an assumption, creating and identifying new elements, etc., see Fig. 1). Since the former research of FMC shares many similarities, the aim of this study is not to construct a new framework but to outline a framework describing how knowledge, practice, and products are connected. On the other hand, this framework fully presents the structure of modeling competence and guides empirical research to evaluate and analyse how modellers apply their knowledge and practices in the process of developing a modeling product. The following paragraph will describe each element and its composition in more detail.

Meta-modelling knowledge was defined as knowledge about models and modelling on "how models are used, why they are used, and what their strengths and limitations are" (Schwarz et al., 2009). Modellers with meta-modelling knowledge could understand the nature of science and reflect on their ability to use and develop scientific models to appreciate how science works and the dynamic nature of knowledge (Abd-El-Khalick et al., 2004; Schwarz & White, 2005). This component is usually introduced to assess learners' knowledge about models and modelling and was classified into the following four dimensions.

- *Nature of models* refers to a representation of a process or phenomenon with a specific purpose; the model is changeable (Lee et al., 2017).
- *Types of models* generally includes six types of models: concrete model, verbal model, visual model, mathematical model, gestural model and mixed model (Boulter & Buckley, 2000).

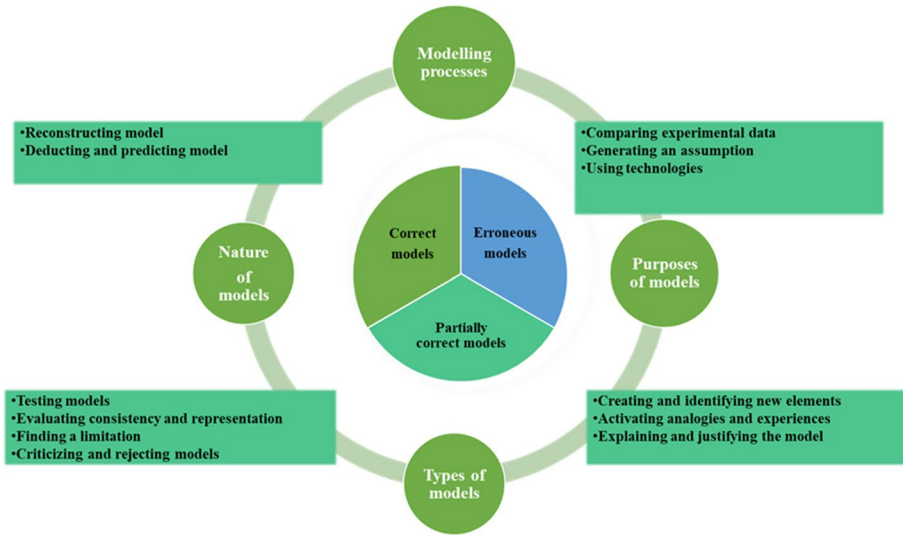


Fig. 1 The framework of modelling competence

- *Purpose of models* includes two categories, basic functions (describing, visualizing and explaining), and advanced functions (standard reference, reasoning, problem solving, communicating, predicting, simulating and generating new ideas) (Lin, 2014).
- *Modelling process* occurs in modelling-based learning or teaching to achieve the educational goal, the generation of models, the evaluation of models and the modification of models (Khan, 2007).

Modelling practices can be defined as cognitive, discursive and social activities that take place in science classrooms, that are related to the modelling process and are aimed at developing epistemic understanding of science concepts and appreciation of the nature of science (Fretz et al., 2002; Jimenez-Liso et al., 2021; Ke & Schwarz, 2019). Science education literature addresses modelling practice in several theoretical frameworks that use overlapping terms to describe its activities, phrases, or processes (Göhner et al., 2022). A number of core modelling practices take place during modelling processes. Louca and Zacharia (2012) found a consensus in the literature about the modelling process involving a stimulus and four discrete steps. Modelling practices are incorporated in the model-based teaching and learning environment, which relates to the theoretical framework used for the description of modelling processes and operationalisation of assessment of modelling processes (Göhner et al., 2022). Since this study aims to explore how participants apply their modelling practices between the experimental world and the modelling world, modelling practices/activities were identified in each discrete step based on research by Göhner and Krell (2022) and Krell et al. (2019) in relation to the BB experiment. This scientific modelling aims to gain understating of a complex real-world system by using relevant modelling activities, which is different from most modelling frameworks as it explores the experimental and physical world rather than mental model development.

Modelling product can be well defined as “the main outcome of any modelling process is the development of a tangible, visible, and communicable artefact that demonstrates the modeller’s understanding and that can be evaluated by specific criteria for

its quality” (Göhner et al., 2022). Modelling products can be used to externalise and express learners’ thoughts (mental models) and help them visualise and examine components of their theory (conceptual models and scientific models) (Gentner & Stevens, 2014; Göhner et al., 2022; Greca & Moreira, 2000; Hamza et al., 2008; Nicolaou & Constantinou, 2014). Regarding the aspect of evaluation in science education literature, modelling products are considered a more content-related approach focusing on the integration of specific components (Baumfalk et al., 2019; Göhner et al., 2022). The modelling product in this study could be classified as correct models, partially correct models, and erroneous models. The criterion of whether the modelling product is good or not depends on the degree to which the modelling products correctly and fully represent the characteristics of the phenomenon; provide a mechanism that accounts for how the phenomenon operates, and are used to formulate predictions about the observable aspects of the phenomenon (Pluta et al., 2011).

## Methods

### Participants

The participants in this study included two groups selected via convenience sampling. These participants were chosen because they shared two characteristics: (1) all of them were enrolled in the same teacher education programme at the same university, chemistry education teacher education programme (CE-TE). (2) they had never experienced formal modelling-based training. This study consisted of 38 pre-service teachers (PSTs) who are studying at a comprehensive public university in the south of China. 38 in-service teachers (ISTs) participated who were all employed at urban schools from primary school to senior high school as science or chemistry teachers. Table 1 contains a summary of these participants’ background information.

Participants in the PST group including seven males and thirty-one females, are studying as Master students (year one), majoring chemistry education, enrolling in the second semester of the first year of CE-TE programme. They are predominantly female, 21–23 years of age (the majority of participants are female, which would be a potential limitation and will be addressed in Limitations section); 92% of PSTs majored in chemistry for their bachelor’s degree, while only three pursued non-chemistry majors, such as business and accounting. However, all of them passed the national master entrance exam, which indicated that all PSTs had the required knowledge and skills in the teacher education programme.

In the context of this study, ISTs ( $n=33$ ) refer to specialist science teachers who teach in primary schools or middle schools, covering grades K1-K9. In China, science is treated as an independent subject with dedicated teachers responsible for its instruction. The teachers involved in this study who work in high schools ( $n=6$ ), all teach chemistry. This is because they all obtained master’s degrees in chemistry education. All of ISTs graduated from the CE-TE programme, and the years of total teaching experience and science teaching experience of participants ranged from one to three years. Araujo et al. (2016) pointed out that teachers with no experience to three years of experience are referred to as “rookies”. Thus, the selected sample of ISTs and ISTs aims to compare if teaching experience could be an influential factor in the modelling competence.

**Table 1** Background information of pre-service teachers and in-service teachers

	Pre-service teachers' group	In-service teachers' group
Average Age	22	25
Gender		
<i>Male</i>	7	6
<i>Female</i>	31	32
Subject major in bachelor's degree		
<i>Chemistry majors</i>	35	38
<i>Non-Chemistry majors</i>	3	0
Teaching level		
<i>Primary school (K1-K6)</i>	/	13
<i>Middle school (K7-K9)</i>	/	20
<i>High school (K9-K12)</i>	/	5
Teaching subject		
<i>Science</i>	/	32
<i>Chemistry</i>	/	6
Teaching experience		
<i>One year</i>	/	19
<i>Two years</i>	/	14
<i>Three years</i>	/	5

## Instruments

Two instruments were developed to measure three components of modelling competence. A Likert scale questionnaire was used to evaluate meta-modelling knowledge and a Black Box (BB) modelling task was applied to assess modelling practices and modelling products. Two different techniques, think aloud method and drawing were used to collect modelling practices and products respectively.

### Meta-modelling knowledge questionnaire

The questionnaire comprised two parts. In the first part, the participants were asked to provide their background information including their personal identifier, their age, gender, major in bachelor's degree, and their total years of teaching experience. The second part consisted of 20 items to collect information about their knowledge of four meta-modelling knowledge dimensions: nature of the model (4 items), the purpose of models (6 items), the types of models (6 items) and the process of modelling (4 items). Each item had five options, strongly disagree, disagree, not sure, agree, and strongly agree, which were assigned 1 to 5 scores respectively. Since the adjective labels on the scale with a meaningful numerical interpretation was used, it is appropriate to treat the rating scale data as interval data rather than ordinal data (Taber, 2018).

Before generating the final questionnaire, the original questionnaire with a total of 35 items had a test process to ensure sufficient validity and reliability. The design of each item in each dimension was based on the previous studies (Grosslight et al., 1991; Lee, 2018; Lin, 2014; Treagust et al., 2002). A pilot study was implemented to test the appropriateness

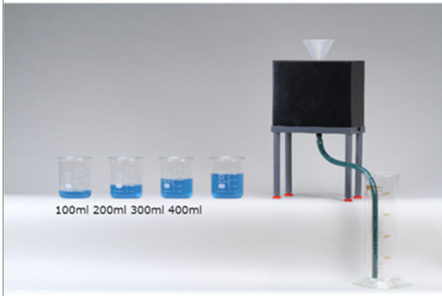
of the instruments for ensuring content validity. Nine participants in the pilot study, including two PSTs who were in year two in the CE-TE programme, two PhD students in science education, two ISTs with one-year teaching science experience in a primary school and one lecturer with PhD degree in science education who is a member of academic staff in the cohort of science education in the CE-TE programme. This helped strengthen the match between the items and their relevant dimensions and simplify the wordings and the expressions of the questionnaire. Confirmatory factor analysis (CFA) was conducted to evaluate whether the proposed structure (four factors model) fit the modellers' answers ( $n=76$ ). A four-factor model was generated with acceptable goodness-of-fit indices ( $\chi^2/df=1.340$ ,  $GFI=0.955$ ,  $CFI=0.916$ ,  $NNF=0.932$ , and  $RSMEA=0.0067$ ). Finally, Cronbach's alpha value for each dimension ranged from 0.739 to 0.931, indicating that the scale had strong internal consistency (McCrae et al., 2011). These indicators showed that the questionnaire could be regarded as reliable.

### Black Box modelling task

To determine whether participants could use modelling practices to generate a modelling product for problem-solving in specific situations, they were asked to work out one modelling task with three rounds of experimental data, namely the BB (see Fig. 2). This task was revised from previous studies by Göhner et al. (2022) and Krell et al. (2019). According to the given data, three different data sets required participants to generate different models. Given data included the input of pouring water into a black box and observing the output of water in each round. The second and third rounds of data provided opportunities to

The structure inside the black box is unknown. Please **infer the internal structure** of the black box by three rounds of experiments, as shown in the figure below. Each time a different volume of beaker was chosen (100ml, 200ml, 300ml, 400ml), or the same volume of beaker was used repeatedly, which were poured into the funnel above the black box, and the change in scale of the measuring cylinder was observed.

Based on the experimental data of each round, please **draw possible models** to show the structure inside the black box. You can also use words, symbols, etc. to help show what you have in mind as you draw your diagram. **Importantly, please “think aloud” and record your talking.** You have **20 minutes** to complete this task.



**According to given data of input and output, graphically draw a model of the inner system of the black box in each round.**

**(1) The first round:**  
 Input: 400, 400  
 Output: 0, 400

**(2) The second round:**  
 Input: 400, 400, 400, 400  
 Output: 0, 400, 600, 400

**(3) The third round:**  
 Input: 400, 400, 400, 400, 400, 400  
 Output: 0, 400, 600, 400, 0, 1000

Fig. 2 The Black Box modelling task

evaluate and modify the models that participants had created. Participants were asked to reason why the output of water changed for each round, generate hypotheses about what was happening, make predictions, and eventually, attempt to infer the hidden mechanics of the box. Participants were instructed to think aloud while modelling the Black Box to gain insight into their modelling processes. The model that they drew from the third-round data sets were considered as the indicator of quality of their modelling products.

## Data analysis

Both quantitative and qualitative analysis were used in this study. Descriptive statistics on the rating scale questionnaire of meta-modelling knowledge were used to calculate the mean and standard deviation of each item. Since the Kolmogorov–Smirnov test indicated the data did not follow a non-normal distribution ( $D=0.1$ ,  $p=0.02 < 0.05$ ), Spearman's correlation coefficient was examined among three components for the PSTs group and ISTs group, respectively, to observe if there were any correlations among the three components. Wilcoxon rank-sum tests were then carried out to examine the differences in performance in terms of meta-modelling, practice and product between the two groups. The effect size  $r$  was calculated. The interpretation values for  $r$  commonly found in published literature (Göhner et al., 2022; Mulder et al., 2016) are:  $0.10 < 0.3$  (small effect),  $0.30 < 0.5$  (moderate effect) and  $\geq 0.5$  (large effect).

An additional step was to quantify qualitative data. Quantifying qualitative data involves converting the subjective and descriptive information into objective data, such as coding data and rating scale data (Krippendorff, 2018; Miles & Huberman, 1994). Regarding evaluating participants' modelling practices, cognitive processes such as solving problems or thinking abstractly can be uncovered by asking participants to think aloud (Ericsson & Simon, 1998). After the training, the participants were asked individually to describe aloud the process they followed to solve the BB modelling task. Their speaking was recorded with electronic devices and then transcribed for coding. A deductive approach was applied since a theoretical conception of the process of scientific modelling in a BB modelling task is available, referring the studies by Krell et al. (2019). However, the deductive approach was supplemented by an inductive refinement of the modelling practices. A rubric of modelling practices established by previous studies (Cheng et al., 2021; Khan & Krell, 2019) and the revised Bloom's taxonomy (Jensen et al., 2014) was used. As shown in Table 2, four levels were created to assess respondents' modelling practices. Since it was difficult to give a score for each sub-modelling practice when participants outlined a complex system, and their responses were highly circular and iterative, nonetheless the presented scoring rubrics showed cognitive levels after interactions while respondents performed BB.

This study employed the draw-a-picture research technique to assess the participants' quality of model products. Drawings in this study included diagrams, graphs, images, or other visual representations combined with text made by the participants. In contrast, producing pure text or essays is not considered drawing. According to the criteria for good scientific models by Namdar and Shen (2015) and Cheng et al. (2021), modelling products were assessed according to quality, based on the degree to which and the way in which model reflected two features: (1) correctness and completeness of symbolic representations (such as the accuracy of the instrument and proficiency of explanatory texts used in a model); (2) how well the whole modelling product coherently reflected the underlying mechanism of the BB. The scoring rubrics of the study used four levels to cover all aspects of the participants' drawings, as shown in Table 3.



**Table 2** The description and examples of a scoring system for assessing modelling practice

Level	Revised Bloom's taxonomy	Modelling practice	Example
1 (Low level)	Remember Understand	Modellers can <b>generate/select</b> a model, <b>identify</b> the variables/elements, and <b>explain/discuss/describe</b> ideas and concepts based on experience, data	<ul style="list-style-type: none"> <li>• <i>I thought of a water outlet pipe at the 400ml position, and once the water exceeded 400ml, the water would overflow (Code: generate a model)</i></li> </ul>
2 (Middle level)	Apply Analyze	Modellers' practice only focuses on <b>one to two relevant aspects: use</b> the information in a new way, <b>interpret</b> the variables /elements, <b>compare</b> different sets of data, <b>distinguish</b> between the different parts of a system, and <b>criticize/question</b> the previous model	<ul style="list-style-type: none"> <li>• <i>When I found that the beaker of water poured 200ml into the black box, the previous BB model was obviously not working (Code: use information and criticize the model)</i></li> </ul>
3 (Middle level)		Modellers' practice only focuses on more than two relevant aspects, but they are treated independently and additively.: <b>use</b> the information in a new way, <b>interpret</b> the variables /elements, <b>compare</b> different sets of data, <b>distinguish</b> between the different parts of a system, and <b>criticize/question</b> the previous model	<ul style="list-style-type: none"> <li>• <i>Compare: In the third round of experiments, I found a big difference, that is, when 400ml of water is input for the fifth time, the output of water is the only 0ml. Then I thought of dividing it into two parts, dividing it into two structures on the left and right, and then one structure on the left and one structure on the right (Code: compare, interpret and distinguish the elements of BB)</i></li> </ul>
4 (High level)	Evaluate Create	Modellers' practice not only contains all aspects in the other three levels but also can <b>justify</b> a stand and <b>create</b> a new model, <b>evaluate, judge, appraise and develop</b> the model	<ul style="list-style-type: none"> <li>• <i>When the water reaches 400ml or more, BB will open the valve, and then when the water level is 200ml, BB will close the valve, so that the output of the water is 200ml</i></li> </ul>



**Table 4** Results of Spearman’s correlation coefficient and significance in different sample settings

		1	2	3
Entire sample (n = 64)	<b>1 meta-modelling knowledge</b>	/	0.08	0.06
	<b>2 modelling practice</b>		/	0.62**
	<b>3 modelling products</b>			/
PST group only (n = 29)	<b>1 meta-modelling knowledge</b>	/	0.22	0.17
	<b>2 modelling practice</b>		/	0.78**
	<b>3 modelling products</b>			/
IST group only (n = 35)	<b>1 meta-modelling knowledge</b>	/	-0.16	0.02
	<b>2 modelling practice</b>		/	0.42*
	<b>3 modelling products</b>			/

\*\*Correlation is significant ( $p < 0.01$ ). \*Correlation is significant ( $p < 0.05$ )

**Table 5** Wilcoxon rank-sum test for the two groups’ three components of modelling competence

Components of modelling competence	PST group		IST group		Z	p	r
	M	SD	M	SD			
Meta-modelling knowledge	3.94	0.40	4.16	0.43	-2.38*	0.02	0.39
Modelling practice	2.45	1.33	2.20	1.02	-0.66	0.51	0.11
Modelling product	2.45	1.02	2.69	0.63	-0.96	0.34	0.17

M: means; SD: standard deviation; r: effect size; \*:  $p < 0.05$

presented in Table 6. The percentages indicate the ratio of the number of a single code to all codes in the group, and the numbers indicate how many times a single code appeared in the group from all responses.

In general, both groups covered all codes of modelling practice, and the frequency of each code had similar percentages. It indicated that both groups performed equally well in terms of modelling practice when solving a complex problem, which was consistent with the quantitative results. The highest frequencies code was *comparing data* for both groups (23% for PSTs, and 31% for ISTs). The frequency of *explaining and justifying the model* also had a high percentage in both groups (19% for PSTs, and 20% for ISTs). It was worth noting that this practice often came together with *creating and identifying the new element*. This result illustrated that even though these novel modellers experienced no or little modelling training, they were able to interpret the constructed models or elements. For the code of *using technologies*, most modellers created a model with physical structures, including pipes, switches, and beakers. Only a few ( $n = 3$ ) used technology, such as using sensors and controlling by programming systems. Taken together, these results suggested that these novice modellers had a variety of modelling processes to solve a complex system, and the frequencies of these modelling processes were nearly identical for both groups.

Figure 3 shows that PSTs’ modelling practices seemed to fall into two groups – the very naïve and the relatively competent (55% level 3 and level 4 in total). However, most ISTs were relatively naïve. There were 43% of PSTs’ modelling practices only at Level 1 with a low level of cognition (13 out of 33). At this level, modellers in both groups usually were able to answer the first round of the BB experiment by drawing a required model and

**Table 6** Frequencies of codes for modelling practice in both groups

Codes for modelling practice	PST group		IST group	
	Numbers	Percentages	Numbers	Percentages
Comparing experimental data	56	23%	78	31%
Activating analogies and experiences	13	5%	18	7%
Creating and identifying new elements	43	18%	47	18%
Evaluating consistency and representation	19	8%	18	7%
Finding a limitation	8	3%	6	2%
Deducting and predicting model	10	4%	5	2%
Generating an assumption	12	5%	8	3%
Criticizing and rejecting models	8	3%	9	4%
Testing models	11	5%	7	3%
Explaining and justifying the model	46	19%	50	20%
Reconstructing models	12	5%	7	3%
Using technologies	2	1%	1	1%

clearly explaining how the water flows through the specific structure. However, they didn't have any idea about giving an answer to the second round of experiments. The percentage of PSTs (31%) at level 4 ( $n=10$ ) was higher than ISTs (14%), who held higher-level aspects of modelling practice, synthesizing different modelling aspects to solve a problem, such as evaluating and reconstructing models ( $n=5$ ). Modellers in this level usually start working on the third round of experiment. By considering more complex conditions that happened from input and output data, the modelling aspects of comparing, evaluating, testing, and revising were iterated through the process.

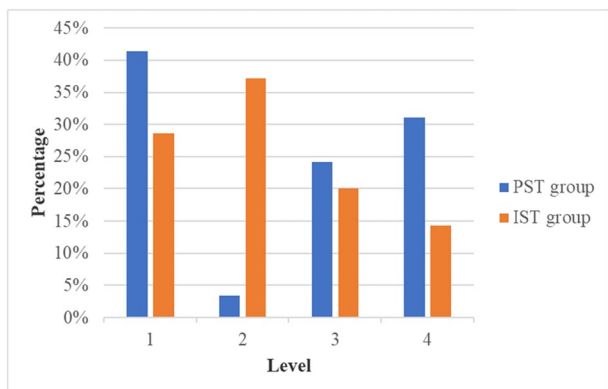
## Modelling products

Drawings were used as the basis for assessing participants' modelling products. In this approach, two groups of teachers created drawings of scientific phenomena (BB) and used the language of icons representing system behaviours and visualizing the modelling process according to each round of experimental data. The final drawing was considered to examine the accuracy and quality of the modelling product. Figure 4 illustrates the percentage of each scoring level of students' drawing for a BB structure in both groups.

Across both conditions, there were relatively few people with level 1 and level 4. There were 21% of PSTs at level 1 (7 out of 33), while less than 3% of IST was at this level (1 out of 35). A relatively small percentage of modellers provided a perfectly correct model (level 4) which could be applicable to the given experimental data. The majority of in-service teachers (60%,  $n=21$ ) developed a partial product with medium quality (level 3). They were able to create a model which basically satisfied the given situation, but a few elements and structure were unclear. Even though 42% ( $n=14$ ) of the PSTs had reached level 3, which was the highest percentage among the four levels, the number was still less than the ISTs. Regarding level 2, 31% of ISTs ( $n=11$ ) and 24% ( $n=8$ ) of PSTs created a partial product with low quality. The low quality of BB appropriately reflected the real phenomenon, but the mechanism was not achievable.

It was interesting to analyse the correct model products and find the constraints of the partially correct model products. A total of six modellers created the correct model products (PST=4, IST=2). Regarding these correct models, there were four modellers (PST=3 and IST=1) who created a physical model product to present the structure of

**Fig. 3** The percentage of each level of modelling practice in both groups



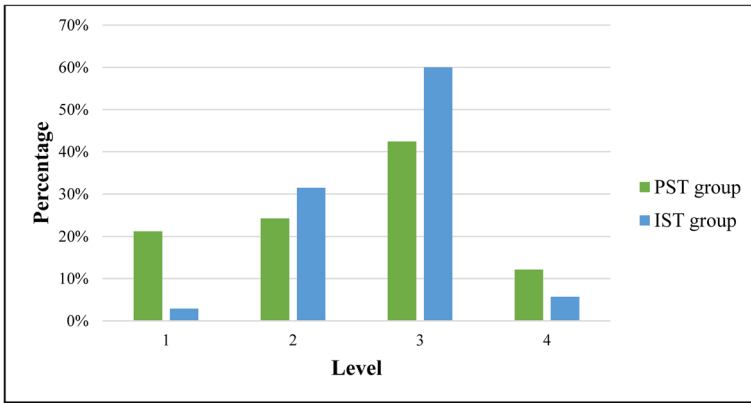
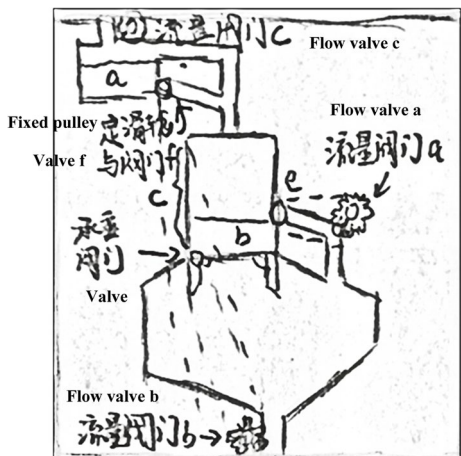


Fig. 4 The percentage of each scoring level of students' drawing for a BB structure in both groups

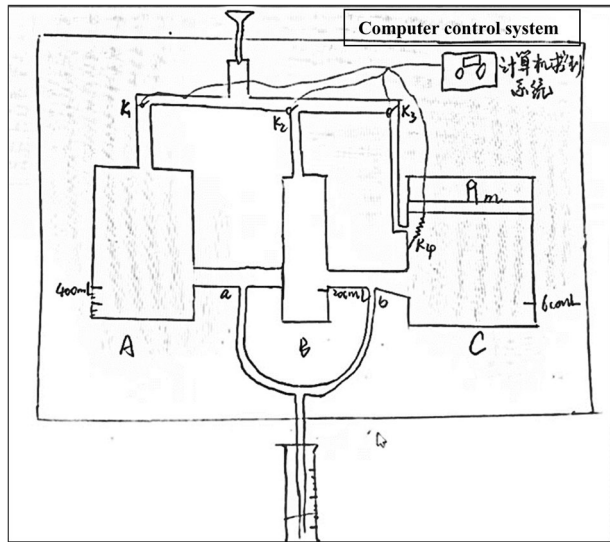
BB. These models included some physical elements, such as pipes, valves, bulkheads, pulleys, and containers with different volumes. Although these products were not consistent with the sample model (presented in the methodology section), they could conform to the experimental data correctly. The other two modellers (PST=1 and IST=1) drew a model by using technological elements. These technologies included sensors and computer-controlled systems. Figures 5 and 6 present the correct examples of a physical model and a technology-supported model respectively.

Figure 5 displays a correct drawing of a physical model (a and b represent containers; c, e and f valves; the dotted line is the connection line of the fixed pulley), which was given by a PST. He illustrated the mechanism of BB: The volume of both containers (a) and (b) was 200 ml. Container (c) was 600 ml. When there were 400 ml of liquid in the system, the gear drove the fixed pulley to close the switch (e). When there were 1400 ml of liquid in the system, the gear drove the fixed pulley to open (f) but close the load-bearing valve. When the volume was 600 ml, the valve was opened. When there was 2400 ml liquid in the system, valve (f) opened. From these detailed descriptions, he succeeded in identifying the relationship between the variables underlying the experimental data and physically

Fig. 5 The correct drawing of a physical model



**Fig. 6** The correct drawing of a technology-supported model



connected them with different elements (e.g., switch, gear, and valve). This result showed this modeller had high-level modelling practice and robust scientific knowledge. The other cases of correct models came with the same results.

Physically creating a model for BB appears to be more challenging, not simply for PSTs or ISTs. An interesting drawing from a PST for creating a correct model included using a computer-supported system to control all elements rather than using complex physical structures (see Fig. 6, note: A, B, and C represent containers; a and b represent pipelines; k1-k4 represent electric valves). To illustrate this product, four switches were controlled by a computer. The computer system could control different switches according to the amount of water intake. Then the water flowed into different paths and finally into the measuring cylinder. Although the technology supported the implementation of the system, it still required design thinking as well as a relevant physical understanding for making it.

## Discussion

In order to address the first research question, this study investigated the relationship among three components of modelling competence in different groups (PST group and IST group). In line with the recent assumption by Göhner et al. (2022), the results further supported the view that modellers with higher modelling practices would develop more qualified modelling products. Inconsistent with the assumptions of meta-modelling knowledge guiding practices (Cheng & Lin, 2015; Schwarz et al., 2009), the analysis of think-aloud data found modellers did not express much related meta-modelling knowledge. In a modelling environment, more important aspects were identified in that modellers applied iterative modelling practices to solve the BB modelling tasks. A product after more modelling activities would have more elements and more explanatory power.

Meta-modelling knowledge, modelling practices and modelling products have distinct contributions to the process of problem-solving. Although meta-modelling information does not affect modelling practices or modelling products, it does offer a fundamental

theoretical framework for comprehending the interactions between different system components and informing modelling activities (Ke & Schwarz, 2016; Schwarz & White, 2005). On the other hand, modelling practises are the processes via which modelling products are developed and enhanced. Modelling practices are techniques that result in high-quality modelling products, hence modelling practices are positively correlated with modelling products. Therefore, modelling products are the outcomes of modelling practices, and their quality is reflected in the accuracy of these results.

By comparing significant differences in responding to meta-modelling knowledge items between the two groups, it was found that ISTs had a higher level of knowledge than PSTs. This finding may indicate that ISTs who had never experienced modelling-based training could advance their meta-modelling knowledge by reflecting on their daily teaching experience. Few studies examined differences in the perspective of meta-modelling between PSTs and ISTs. The presented results were aligned with the studies of Justi and Gilbert (2002a, 2002b, 2003), which indicated experienced teachers had awareness of the value of models in learning science. However, Van Driel and Verloop (2002) investigated modelling knowledge in different groups of experienced teachers who had different science backgrounds and found that teachers' subject and teaching experience had no relationship with their responses towards models and modelling in science. This inconsistency may be due to the different national curricula and contexts (Gogolin & Krüger, 2018). Informal teaching activities, such as learning from teaching resources and peer learning, occurring in ISTs' daily school environment were able to help them engage in various model and modelling scenarios.

Although the study found that meta-modelling knowledge appeared to be related to teachers' actual teaching experience (with significant differences in performance between the two groups), there was no difference in the performance of the two groups in terms of modelling practices. Looking further, the levels of modelling practices in the PST group were divided into two main groups, naïve and relatively competent. In contrast, ISTs mainly performed lower-level practices. Consistent with previous studies, teachers with different teaching experience teaching science subjects did not differ significantly in their modelling performance, unless they were experts (Hogan et al., 2003; Kang et al., 2018; Van Driel & Verloop, 2002). In the actual teaching, ISTs' modelling practices are mainly reflected in their instructional sequences, that is, how teachers apply model pedagogy to support students to construct scientific knowledge (Schnotz & Bannert, 2003; Tytler & Hubber, 2016; Van Driel & Verloop, 1999; Xue & Sun, 2022). Actually, they rarely have the opportunity to experience authentic modelling scenarios (Gilbert, 2004; Prins et al., 2009, 2011). Therefore, support and professional learning are a priority for all teachers, not just beginners, and providing an authentic modelling environment is needed to promote teachers' modelling practices (Graham et al., 2020; Hamza et al., 2008; Stammen et al., 2018).

An innovative aspect of the study is the analysis of the modelling product, which has received little attention in previous related studies, as stated by Chiu and Lin (2019) and Göhner et al. (2022). The results revealed that both groups were able only to develop a very limited number of correct models, but more partially correct models. Literature suggests modelling products as indicators for evaluating students' meta-modelling knowledge and modelling practices (Cheng et al., 2021; Göhner et al., 2022; Krell et al., 2019; Schwarz et al., 2007). The presented results indicated that modellers with sophisticated meta-modelling knowledge would not necessarily create high-quality modelling products. The reason might be that a correct modelling product needs not only high-level modelling competence but also scientific knowledge and other skills, such as hands-on skills, technical literacy and drawing (Göhner et al., 2022). Regarding the BB modelling task in this study, the scientific knowledge of water pressure and siphoning would influence



participants to develop their modelling products by observing the data from the amount of input water and output water. In addition, some participants designed technology-based elements in their drawing (sensors and artificial intelligence) to describe the inside structure of BB rather than the physical structure that most participants created.

## Limitations

One of the main limitations of this study concerns sampling. Participants in the present study were not randomly assigned to groups (Ledford, 2018). The researcher was only permitted to gather data at the selected university, specifically in the chemistry cohort. The study sample consisted solely of females, which may limit the generalizability of the results to male populations or populations with a mix of males and females. It is important to consider gender diversity in future studies to ensure that the findings are valid across different genders. Regarding ISTs, a few were from high schools and most of them came from K1-K9. There were likely differences in teacher training, pedagogy and scientific skills, which may have some influence on the generalisability of the result. On another hand, the selection of only ISTs with 1–3 years of experience may limit the generalizability of the findings to all in-service teachers, including those with more or less teaching experience. It is important to recognize this sample limitation when interpreting and applying the results of this study to other populations of teachers. Future studies could consider including teacher candidates with a wider range of experience to enhance the generalizability of the findings. The study's results would be stronger if it used random sampling, as it ensures that results obtained from the sample could represent the entire level of science teacher education program as well as presenting more comprehensive picture how ISTs across different educational level perform modelling competence. Another limitation would be the Black Box approach. Regarding the assessment of the modeling processes, the black box as a modeling task may have limited the participants in their engagement, as it is a rather abstract and complex task (Leden et al., 2020). Although the developed BB task was simplified based on the original one, some instructions were still needed before modellers start to answer the task in case all participants could fully understand. Finally, using proposed a scoring system for the overall analysis of modelling practice, the researcher could monitor and assess the overall performance of modelling practice. However, it is acknowledged that this scoring system is a somewhat nebulous way to describe the entire modelling process. It is impossible to say that a modeller with level 4 advanced modelling skills is competent enough to use all of the modelling activities in this research. Therefore, in using the defined scoring criteria for modelling practices, researchers may evaluate the modelling activities that modellers use and how they integrate them (independently or additively).

## Conclusions and implication

Modelling competence has been characterised by different studies, which indicated modellers with sufficient model and modelling knowledge conducted different modelling practices to generate modelling products to respond to specific questions both theoretically and in authentic scenarios (Weinert, 2001). The presented research proposed three components of the framework of modelling competence (meta-modelling knowledge, modelling practice and modelling products), and investigated the two groups'

(PSTs and ISTs) relationships between the three components. This aimed to enrich the development of theory and practice in modelling competence, as well as fill research gaps in the assessments of modelling competence in science teacher education.

To respond to the first research question, this study showed that meta-modelling knowledge has no relationship with either modelling practices or modelling products, but modelling practices were positively correlated with modelling products. This implies that during the process of problem-solving, meta-modelling knowledge did not play a measurable role in guiding modelling practice and generating high-quality modelling products. Modelling practice at a higher level was reflected in the analysis of correct model products. Regarding the second research question, both groups were at a high level of meta-modelling knowledge, but the ISTs were higher than the PSTs. The two groups did not perform well in modelling practices and products and there was no difference between the two groups. A high modelling product not only requires the modeller to apply high level modelling practices, but also more practical abilities, which may be closely related to STEAM disciplines.

The findings are instructive in that in-service teachers' experience of teaching in schools was helpful in improving their meta-modelling knowledge, but their modelling practices and products were the same as those of ISTs, which is informative for future IST professional development. The results also contribute to filling the gap in the literature on modelling competence development in pre-service and in-service professional development (Göhner et al., 2022; Nielsen & Nielsen, 2021). Another contribution is that the development and validation of modelling competence assessment tools and scoring rubrics have been proposed for process-oriented assessment rather than product-oriented assessment. The BB modelling task is often used in science education to assess students' scientific literacy (Göhner & Krell, 2022; Göhner et al., 2022; Krell et al., 2019). Previous research has used physical black boxes to assess students' modelling practices, which allows students to experience real-life scenarios but is not suitable for large-scale testing (it is usually for case studies). This study adapted a dynamic task to a static one and fitted the characteristics of the modelling task, i.e., a description of the stem, followed by different experimental data to guide students in generating a model, evaluating it and modifying it. Students' modelling practice was analysed through a think-aloud approach and the modelling products were assessed in the form of drawings. The findings suggested that the task and scoring rubrics developed can be used to assess modelling practices and products. The process of assessment went through a pilot study and formal experimental testing, and this work strengthened the reliability and validity of the research instrument.

This study equips the global science education community with a methodical framework related to modelling competence that can be employed in diverse educational contexts, thereby amplifying its relevance to the primary or middle school level. For example, the research model can serve as a robust assessment tool for evaluating primary and middle school students' abilities in scenario-based modelling tasks, providing educators with insights into areas of strength and needed improvement. Secondly, the research model could be considered as a theoretical foundation upon which curricula can be designed, specifically aiming to bolster the facilitation and understanding of modelling competence, ensuring students not only grasp the concept but also proficiently apply it in varied scientific contexts. Regarding the fact of research outcomes, it emphasizes a universal area in need of improvement in science teacher education, regardless of teaching experience. While theoretical knowledge is essential, its translation to practical application remains challenging. This distinction is vital for shaping the curriculum

in teacher training programs, ensuring that educators are not just theoretically equipped but also adept at practical applications. By understanding these disparities and their implications, this study paves the way for targeted professional development programs, ensuring teachers are better prepared to facilitate the intricacies of modelling competence to their students, especially in primary and middle years.

Future studies could continue to develop the theoretical framework with the three components and incorporate it into different educational levels across disciplines and modelling tasks. To respond to the latest call by Schwarz et al. (2022) in efforts to support and assess the practices, it is important for further studies to explore and assess individual modelling activities, identifying which activity influences overall modelling practices and what impact it has on the final products. Finally, modelling practices have a positive relationship with modelling products, while qualitative data indicated that the scientific knowledge held by the modeller, as well as engineering skills, affect modelling practices as well as products. More evidence is needed to support this view.

## Declarations

**Ethical approval** The research in this study was conducted with approval from the University of Dundee, School of Education and Social Work Research Ethics Committee (E2020-66).

**Conflict of interest** The authors declare no competing interests.

## References

- Abd-El-Khalick, F., Boujaoude, S., Duschl, R., Lederman, N. G., Mamlok-Naaman, R., Hofstein, A., Niaz, M., Treagust, D., & Tuan, H. (2004). Inquiry in science education: International perspectives. *Science Education*, 88(3), 397–419. <https://doi.org/10.1002/sce.10118>
- Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2016). Teacher quality and learning outcomes in kindergarten. *The Quarterly Journal of Economics*, 131(3), 1415–1453. <https://doi.org/10.1093/qje/qjw016>
- Baumfalk, B., Bhattacharya, D., Vo, T., Forbes, C., Zangori, L., & Schwarz, C. (2019). Impact of model-based science curriculum and instruction on elementary students' explanations for the hydrosphere. *Journal of Research in Science Teaching*, 56(5), 570–597. <https://doi.org/10.1002/tea.21514>
- Boulter, C. J., & Buckley, B. C. (2000). Constructing a typology of models for science education. *Developing Models in Science Education*, 41–57. Springer, Dordrecht. [https://doi.org/10.1007/978-94-010-0876-1\\_3](https://doi.org/10.1007/978-94-010-0876-1_3)
- Cheng, M. F., & Lin, J. L. (2015). Investigating the relationship between students' views of scientific models and their development of models. *International Journal of Science Education*, 37(15), 2453–2475. <https://doi.org/10.1080/09500693.2015.1082671>
- Cheng, M. F., Wu, T. Y., & Lin, S. F. (2021). Investigating the relationship between views of scientific models and modeling practice. *Research in Science Education*, 51(1), 307–323. <https://doi.org/10.1007/s11165-019-09880-2>
- Chiu, M.H., Lin, J.W. (2022). Research on modeling competence in science education from 1991 to 2020 with cultural and global implications. In: Atwater, M.M. (eds) International Handbook of Research on Multicultural Science Education. Springer International Handbooks of Education. Springer, Cham. [https://doi.org/10.1007/978-3-030-83122-6\\_34](https://doi.org/10.1007/978-3-030-83122-6_34)
- Chiu, M.-H., & Lin, J.-W. (2019). Modeling competence in science education. *Disciplinary and Interdisciplinary Science Education Research*, 1(1), 1–11. <https://doi.org/10.1186/s43031-019-0012-y>
- Danusso, L., Testa, I., & Vicentini, M. (2010). Improving prospective teachers' knowledge about scientific models and modelling: Design and evaluation of a teacher education intervention. *International Journal of Science Education*, 32(7), 871–905. <https://doi.org/10.1080/09500690902833221>

- Ericsson, K. A., & Simon, H. A. (1998). How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture, and Activity*, 5(3), 178–186.
- Fretz, E. B., Wu, H.-K., Zhang, B., Davis, E. A., Krajcik, J. S., & Soloway, E. (2002). An investigation of software scaffolds supporting modeling practices. *Research in Science Education*, 32(4), 567–589. <https://doi.org/10.1023/A:1022400817926>
- Gentner, D., & Stevens, A. L. (2014). *Mental Models*. Psychology Press.
- Gilbert, J. K. (2004). Models and modelling: Routes to more authentic science education. *International Journal of Science and Mathematics Education*, 2(2), 115–130. <https://doi.org/10.1007/s10763-004-3186-4>
- Gilbert, J. K., & Justi, R. (2016). Approaches to modelling-based teaching. In *Modelling-based Teaching in Science Education. Models and Modeling in Science Education*, 9. Springer, Cham. [https://doi.org/10.1007/978-3-319-29039-3\\_4](https://doi.org/10.1007/978-3-319-29039-3_4)
- Gogolin, S., & Krüger, D. (2018). Students' understanding of the nature and purpose of models. *Journal of Research in Science Teaching*, 55(9), 1313–1338. <https://doi.org/10.1002/tea.21453>
- Göhner, M. F., Bielik, T., & Krell, M. (2022). Investigating the dimensions of modeling competence among preservice science teachers: Meta-modeling knowledge, modeling practice, and modeling product. *Journal of Research in Science Teaching*, 59(8), 1354–1387. <https://doi.org/10.1002/tea.21759>
- Göhner, M., & Krell, M. (2022). Preservice science teachers' strategies in scientific reasoning: The Case of Modeling. *Research in Science Education*, 52(2), 395–414. <https://doi.org/10.1007/s11165-020-09945-7>
- Graham, L. J., White, S. L., Cologon, K., & Pianta, R. C. (2020). Do teachers' years of experience make a difference in the quality of teaching? *Teaching and Teacher Education*, 96, 103190. <https://doi.org/10.1016/j.tate.2020.103190>
- Greca, I. M., & Moreira, M. A. (2000). Mental models, conceptual models, and modelling. *International Journal of Science Education*, 22(1), 11. <https://doi.org/10.1080/095006900289976>
- Grosslight, L., Unger, C., Jay, E., & Smith, C. L. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching*, 28(9), 799–822. <https://doi.org/10.1002/tea.3660280907>
- Hamza, K., Wickman, P.-O., & Kelly, G. (2008). Using a guided inquiry and modeling instructional framework (EIMA) to support preservice K-8 science teaching. *Science Education*, 92(1), 141–164. <https://doi.org/10.1002/sce>
- Hogan, T., Rabinowitz, M., & Craven, J. A., III. (2003). Representation in teaching: Inferences from research of expert and novice teachers. *Educational Psychologist*, 38(4), 235–247. [https://doi.org/10.1207/S15326985EP3804\\_3](https://doi.org/10.1207/S15326985EP3804_3)
- Jensen, J. L., McDaniel, M. A., Woodard, S. M., & Kummer, T. A. (2014). Teaching to the test... or testing to teach: Exams requiring higher order thinking skills encourage greater conceptual understanding. *Educational Psychology Review*, 26(2), 307–329. <https://doi.org/10.1007/s10648-013-9248-9>
- Jimenez-Liso, M. R., Martinez-Chico, M., Avraamidou, L., & López-Gay Lucio-Vallegas, R. (2021). Scientific practices in teacher education: The interplay of sense, sensors, and emotions. *Research in Science & Technological Education*, 39(1), 44–67. <https://doi.org/10.1080/02635143.2019.1647158>
- Justi, R. S., & Gilbert, J. K. (2002a). Modelling, teachers' views on the nature of modelling, and implications for the education of modellers. *International Journal of Science Education*, 24(4), 369–387. <https://doi.org/10.1080/09500690110110142>
- Justi, R. S., & Gilbert, J. K. (2002b). Science teachers' knowledge about and attitudes towards the use of models and modelling in learning science. *International Journal of Science Education*, 24(12), 1273–1292. <https://doi.org/10.1080/09500690210163198>
- Justi, R. S., & Gilbert, J. K. (2003). Teachers' views on the nature of models. *International Journal of Science Education*, 25(11), 1369–1386. <https://doi.org/10.1080/0950069032000070324>
- Kang, E. J. S., Donovan, C., & McCarthy, M. J. (2018). Exploring elementary teachers' pedagogical content knowledge and confidence in implementing the NGSS science and engineering practices. *Journal of Science Teacher Education*, 29(1), 9–29. <https://doi.org/10.1080/1046560X.2017.1415616>
- Ke, L., & Schwarz, C. V. (2016). Examining the influence of teacher's framing of modeling practices on elementary students' engagement in modeling. *Proceedings of International Conference of the Learning Sciences, ICLS*, 2, 803–806. <https://doi.org/10.22318/icls2016.107>
- Ke, L., Schwarz, C. V. (2019). Using epistemic considerations in teaching: Fostering students' meaningful engagement in scientific modeling. In A. Upmeyer zu Belzen, D. Krüger, & J. van Driel (Eds.), *Towards a Competence-Based View on Models and Modeling in Science Education. Models and Modeling in Science Education*, 12. Springer, Cham. [https://doi.org/10.1007/978-3-030-30255-9\\_11](https://doi.org/10.1007/978-3-030-30255-9_11)

- Khan, S. (2007). Model-based inquiries in chemistry. *Science Education*, 91(6), 877–905. <https://doi.org/10.1002/sce.20226>
- Khan, S., & Krell, M. (2019). Scientific reasoning competencies: A case of preservice teacher education. *Canadian Journal of Science, Mathematics and Technology Education*, 19, 446–464. <https://doi.org/10.1007/s42330-019-00063-9>
- Koepfen, K., Hartig, J., Klieme, E., & Leutner, D. (2008). Current issues in competence modeling and assessment. *Journal of Psychology*, 216(2), 61–73. <https://doi.org/10.1027/0044-3409.216.2.61>
- Krell, M., Walzer, C., Hergert, S., & Krüger, D. (2019). Development and application of a category system to describe pre-service science teachers' activities in the process of scientific modelling. *Research in Science Education*, 49(5), 1319–1345. <https://doi.org/10.1007/s11165-017-9657-8>
- Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage publications. <https://doi.org/10.4135/9781071878781>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174.
- Leden, L., Hansson, L., & Ideland, M. (2020). The mangle of school science practice: Teachers' negotiations of two nature of science activities at different levels of contextualization. *Science Education*, 104(1), 5–26. <https://doi.org/10.1002/sce.21553>
- Ledford, J. R. (2018). No randomization? No problem: Experimental control and random assignment in single case research. *American Journal of Evaluation*, 39(1), 71–90. <https://doi.org/10.1177/1098214017723110>
- Lee, S. W. Y. (2018). Identifying the item hierarchy and charting the progression across grade levels: Surveying Taiwanese students' understanding of scientific models and modeling. *International Journal of Science and Mathematics Education*, 16(8), 1409–1430. <https://doi.org/10.1007/s10763-017-9854-y>
- Lee, S. W. Y., Chang, H. Y., & Wu, H. K. (2017). Students' views of scientific models and modeling: Do representational characteristics of models and students' educational levels matter? *Research in Science Education*, 47, 305–328. <https://doi.org/10.1007/s11165-015-9502-x>
- Lin, J. W. (2014). Elementary school teachers' knowledge of model functions and modeling processes: A comparison of science and non-science majors. *International Journal of Science and Mathematics Education*, 12(5), 1197–1220. <https://doi.org/10.1007/s10763-013-9446-4>
- Louca, L. T., & Zacharia, Z. C. (2012). Modeling-based learning in science education: Cognitive, metacognitive, social, material and epistemological contributions. *Educational Review*, 64(4), 471–492. <https://doi.org/10.1080/00131911.2011.628748>
- McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review*, 15(1), 28–50. <https://doi.org/10.1177/1088868310366253>
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. Sage.
- Ministry of Education. (2022). Ministry of Education on the issuance of compulsory education curriculum standards (2022 Edition). [http://www.moe.gov.cn/srcsite/A26/s8001/202204/t20220420\\_619921.html](http://www.moe.gov.cn/srcsite/A26/s8001/202204/t20220420_619921.html). Accessed 17 Sept 2022
- Mulder, Y. G., Bollen, L., De Jong, T., & Lazonder, A. W. (2016). Scaffolding learning by modelling: The effects of partially worked-out models. *Journal of Research in Science Teaching*, 53(3), 502–523. <https://doi.org/10.1002/tea.21260>
- 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 Research Council. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. National Academies Press <https://doi.org/10.17226/13165>
- 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>
- Nielsen, S. S., & Nielsen, J. A. (2021). A competence-oriented approach to models and modelling in lower secondary science education: Practices and rationales among Danish teachers. *Research in Science Education*, 51, 565–593. <https://doi.org/10.1007/s11165-019-09900-1>
- Pluta, W. J., Chinn, C. A., & Duncan, R. G. (2011). Learners' epistemic criteria for good scientific models. *Journal of Research in Science Teaching*, 48(5), 486–511. <https://doi.org/10.1002/tea.20415>
- Prins, G. T., Bulte, A. M. W., Van Driel, J. H., & Pilot, A. (2009). Students' involvement in authentic modelling practices as contexts in chemistry education. *Research in Science Education*, 39(5), 681–700. <https://doi.org/10.1007/s11165-008-9099-4>
- Prins, G. T., Bulte, A. M. W., & Pilot, A. (2011). Evaluation of a design principle for fostering students' epistemological views on models and modelling using authentic practices as contexts for learning in

- chemistry education. *International Journal of Science Education*, 33(11), 1539–1569. <https://doi.org/10.1080/09500693.2010.519405>
- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, 13(2), 141–156. [https://doi.org/10.1016/S0959-4752\(02\)00017-8](https://doi.org/10.1016/S0959-4752(02)00017-8)
- Schwarz, C. V., & White, B. Y. (2005). Metamodeling knowledge: Developing students' understanding of scientific modeling. *Cognition and Instruction*, 23(2), 165–205. [https://doi.org/10.1207/s1532690xc12302\\_1](https://doi.org/10.1207/s1532690xc12302_1)
- Schwarz, C. V., Meyer, J., & Sharma, A. (2007). Technology, pedagogy, and epistemology: Opportunities and challenges of using computer modeling and simulation tools in elementary science methods. *Journal of Science Teacher Education*, 18(2), 243–269. <https://doi.org/10.1007/s10972-007-9039-6>
- Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., Shwartz, Y., Hug, B., & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, 46(6), 632–654. <https://doi.org/10.1002/tea.20311>
- Schwarz, C. v., Ke, L., Salgado, M., & Manz, E. (2022). Beyond assessing knowledge about models and modeling: Moving toward expansive, meaningful, and equitable modeling practice. *Journal of Research in Science Teaching*. <https://doi.org/10.1002/TEA.21770>
- Sins, P. H. M., Savelsbergh, E. R., van Joolingen, W. R., & van Hout-Wolters, B. H. A. M. (2009). The relation between students' epistemological understanding of computer models and their cognitive processing on a modelling task. *International Journal of Science Education*, 31(9), 1205–1229. <https://doi.org/10.1080/09500690802192181>
- Stammen, A., Malone, K., & Irving, K. (2018). Effects of modeling instruction professional development on biology teachers' scientific reasoning skills. *Education Sciences*, 8(3), 119. <https://doi.org/10.3390/educsci8030119>
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48(6), 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Treagust, D. F., Chittleborough, G., & Mamiala, T. L. (2002). Students' understanding of the role of scientific models in learning science. *International Journal of Science Education*, 24(4), 357–368. <https://doi.org/10.1080/09500690110066485>
- Tytler, R., & Hubber, P. (2016). Constructing representations to learn science. In: Hand, B., McDermott, M., Prain, V. (eds) Using multimodal representations to support learning in the science classroom. Springer, Cham. [https://doi.org/10.1007/978-3-319-16450-2\\_9](https://doi.org/10.1007/978-3-319-16450-2_9)
- Upmeyer zu Belzen, A., van Driel, J., & Krüger, D. (2019). Introducing a framework for modeling competence. *Towards a Competence-based View on Models and Modeling in Science Education*. Switzerland: Springer.
- Van Driel, J. H., & Verloop, N. (1999). Teachers' knowledge of models and modelling in science. *International Journal of Science Education*, 21(11), 1141–1153. <https://doi.org/10.1080/095006999290110>
- Van Driel, J. H., & Verloop, N. (2002). Experienced teachers' knowledge of teaching and learning of models and modelling in science education. *International Journal of Science Education*, 24(12), 1255–1272. <https://doi.org/10.1080/09500690210126711>
- Vo, T., Forbes, C. T., Zangori, L., & Schwarz, C. V. (2015). Fostering third-grade students' use of scientific models with the water cycle: Elementary teachers' conceptions and practices. *International Journal of Science Education*, 37(15), 2411–2432. <https://doi.org/10.1080/09500693.2015.1080880>

- Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Eds.), *Defining and Selecting Key Competencies* (pp. 45–65). Hogrefe & Huber Publishers.
- Xue, S., & Sun, D. (2022). Integrating analogy into scientific modeling for students' active learning in chemistry education. *In Active Learning: Research and Practice for STEAM and social sciences education*. London: Tech-Open Access Publisher.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

## Authors and Affiliations

Song Xue<sup>1</sup>  · Keith Topping<sup>2</sup>  · Elizabeth Lakin<sup>2</sup> · Moritz Krell<sup>3</sup> 

✉ Song Xue  
xuesong@zjnu.edu.cn

Keith Topping  
k.j.topping@dundee.ac.uk

Elizabeth Lakin  
e.lakin@dundee.ac.uk

Moritz Krell  
krell@leibniz-ipn.de

<sup>1</sup> College of Education, Zhejiang Normal University, Zhejiang, China

<sup>2</sup> School of Humanities, Social Sciences and Law, University of Dundee, Dundee, UK

<sup>3</sup> IPN-Leibniz Institute for Science and Mathematics Education, Kiel, Germany