Received: 1 February 2021 | Revised: 10 January 2022 | Accepted: 25 January 2022

DOI: 10.1002/tea.21759

RESEARCH ARTICLE

JRST | WILEY

Investigating the dimensions of modeling competence among preservice science teachers: Meta-modeling knowledge, modeling practice, and modeling product

Maximilian Felix Göhner[1] | **Tom Bielik**[1] | **Moritz Krell**[2]

[1]Faculty of Biology Education, Freie Universität Berlin, Berlin, Germany

[2]Department Biology Education, Leibniz Institute for Science and Mathematics Education, Kiel, Germany

Correspondence
Maximilian Felix Göhner, Faculty of Biology Education, Freie Universität Berlin, Berlin, Germany.
Email: maximilian.goehner@fu-berlin.de

Abstract

Worldwide, teachers are expected to engage their students in authentic practices, like scientific modeling. Research suggests that teachers experience challenges when integrating modeling in their classroom instruction, with one explanation that teachers themselves lack the necessary modeling competence. Currently, theoretical conceptualizations structure the modeling competence into three dimensions: meta-modeling knowledge, modeling practice, and modeling products. While each of these dimensions is well researched on its own and the three dimensions are commonly expected to be highly positively related, studies investigating their specific relationships are widely lacking. Aiming to fill this gap, the present study investigated the meta-modeling knowledge, modeling practice, and modeling products of 35 secondary preservice biology teachers engaging in a black box modeling task. Data were collected with an established pen-and-paper questionnaire consisting of five constructed response items assessing meta-modeling knowledge and by videotaping the participants engaging in the black box modeling task. Herein, the three dimensions of modeling competence were operationalized as five variables including decontextualized and contextualized meta-modeling knowledge, complexity, and homogeneity of the modeling processes and a modeling product score. In contrast to our expectations and common assumptions in the literature, significant relationships between the five variables were widely lacking. Only the complexity of the modeling processes correlated significantly with the quality of the modeling products. To investigate this relationship further, a qualitative in-depth analysis of two cases is presented. Implications for biology teacher education are discussed.

KEYWORDS

preservice science teachers, modeling practice, modeling competence, meta-modeling knowledge, modeling product

1 | INTRODUCTION

Influential science education standard documents, such as the Next Generation Science Standards in the United States (NGSS Lead States, 2013), call upon teachers to engage their students in authentic scientific practices (Osborne, 2014). Since modeling stands at the heart of any authentic scientific endeavor (Giere, 1999), it is seen as an essential part of science teaching and learning (Gilbert & Justi, 2016; Upmeier zu Belzen, Krüger, et al., 2019). Consequently, scientific modeling is proposed to stand in the center of science curricula, incorporating other scientific practices as well (Lehrer & Schauble, 2006; Meister et al., 2021; Passmore et al., 2014; Windschitl et al., 2008). However, this focus on scientific practices demands preservice science teachers to develop the necessary prerequisites ("professional competences"; Baumert & Kunter, 2013) needed to be able to plan lessons, teach, and reflect upon the teaching-learning processes professionally. Hence, next to general pedagogical knowledge and competences, preservice science teachers need to develop competences related to scientific practices—such as modeling—as part of their professional competences (Osborne, 2014); with modeling being one of such competences.

Empirical research has shown that teachers' epistemic ideas about models and modeling (i.e., meta-modeling knowledge) impacts their classroom instruction (Harlow et al., 2013; Vo et al., 2015). However, meta-modeling knowledge is only one dimension in the current theoretical conceptualization of modeling competence, which also encompasses the teachers' abilities to engage in the modeling practices (Nicolaou & Constantinou, 2014) and their abilities to develop a high-quality modeling product (Chiu & Lin, 2019). Although, each of those dimensions of modeling competence is addressed by empirical research, the relationships between preservice science teachers' meta-modeling knowledge, their modeling practices, and modeling product are to date rarely studied (Cheng et al., 2021; Chiu & Lin, 2019; Louca & Zacharia, 2012; Nielsen & Nielsen, 2021a) and was consequently emphasized as "[o]ne of the most pressing needs for future research" (Louca & Zacharia, 2012, p. 486).

1.1 | Models and modeling in science education

Models are central tools for communicating and reasoning in science and essential to scientists for explaining phenomena and for predicting possible outcomes (Giere et al., 2006; Godfrey-Smith, 2006; Harré, 1970). Consequently, modeling competence is emphasized in standards and curricula in many countries (ACARA, 2015; BCMOE, 2019; KMK, 2020; NGSS Lead States, 2013; NRC, 2012; VCAA, 2016). Based on theoretical approaches in the philosophy of science, scientific models can be defined as epistemic tools for sense-making (Knuuttila, 2011). Accordingly, scientific modeling is the iterative and cyclic process of developing and using models in science, aiming at investigating, representing, explaining, and predicting phenomena (Giere et al., 2006; Passmore et al., 2014).

1.2 | Science teachers′ competences related to models and modeling in science

One goal of teacher education is to equip future teachers with the prerequisites needed to plan lessons, teach, and reflect upon the teaching-learning processes professionally (Baumert & Kunter, 2013; Carlson & Daehler, 2019). Science teachers need to have meta-modeling knowledge as well as the abilities to engage in modeling practices and to develop high-quality modeling products (Chiu & Lin, 2019; Nicolaou & Constantinou, 2014). Furthermore, science teachers need related pedagogical content knowledge (PCK), including knowledge about teaching with and about models and about how to conduct modeling activities in science classrooms (Justi & Van Driel, 2006).

A multitude of research has been conducted on investigating pre- and in-service science teachers' professional competences related to models and modeling, including, but not limited to, teachers' instructional practice regarding models and modeling in classrooms, as well as their own meta-modeling knowledge and modeling practice (Khan, 2011; Krell & Krüger, 2016; Oh & Oh, 2011; Shi et al., 2021; Torres & Vasconcelos, 2015; Vo et al., 2015, 2019).

Most studies indicate that classroom practice typically gives students few opportunities to meaningfully engage with models (Campbell et al., 2015; Khan, 2011). Studies on the instructional practice of teachers in science classrooms suggest that teachers encounter a multitude of challenges and mainly focus on knowledge aspects represented in models, disregarding the predictive nature of models (Harlow et al., 2013; Nielsen & Nielsen, 2021b; Shi et al., 2021; Vo et al., 2015, 2019). However, it is commonly assumed, that teachers' own ideas about models and modeling shape their instructional classroom practice (Harlow et al., 2013; Vo et al., 2015, 2019). For example, Vo et al. (2015) observed that those teachers, who more strongly emphasized otherwise uncommon epistemic ideas (*evidence*, *mechanism*, and *audience*) and modeling practices (*evaluate* and *revise*), employed better instructions during classroom practice. However a later longitudinal investigation by the same authors, suggests this transfer is likely delayed and takes place over vast timescales, often leading to changes in instructional practice only after a year (Vo et al., 2019).

Keeping in mind that teachers' own epistemic ideas can impact their classroom instruction, it is alarming that most studies focusing on pre- and in-service science teachers' meta-modeling knowledge unravel rather uninformed views of teachers on models and modeling in science, including naïve realist views of models as simple copies of reality (e.g., Krell & Krüger, 2016; Torres & Vasconcelos, 2015). Moreover, only a fraction of science teachers expresses that models, as well as the underlying ideas represented in models, may be tested by deducing predictions (Krell & Krüger, 2016).

Studies investigating teachers' modeling practice are rare, but the existing studies come to similar results. It is, for example, consistently observed that some aspects of modeling practice are challenging for teachers, especially the evaluation of models (Khan, 2011; Vo et al., 2015), and, again, the predictive use of models (Göhner & Krell, 2020a). Göhner and Krell (2020a) investigated the modeling practice of 32 secondary preservice biology teachers investigating a black box. They observed that only 14 of these 32 secondary preservice biology teachers used their developed models to predict the investigated system's behavior. Moreover, even among these 14 secondary preservice biology teachers, the developed models were rarely evaluated based on predictions in a repeated and systematic manner.

In summary, studies addressing teachers' instructional practice regarding modeling their epistemic ideas about models and modeling, and their modeling practice, suggest that the potential of models as epistemic tools in teaching scientific practices is left untapped (Harlow et al., 2013; Nielsen & Nielsen, 2021b). The existing body of empirical research, theoretical considerations, and curricula highlight the epistemic nature of models as tools for scientific reasoning. However, in classroom settings models are still predominantly used by teachers for the purpose of communicating scientific content, who are likely not fully aware of the potential, models and modeling may have. While teachers' epistemic ideas of models and their relation to classroom instruction are quite well researched, studies connecting teachers' modeling practice with their classroom instruction and connecting teachers' epistemic ideas of models with their modeling practice are widely lacking. Our study aims to fill this gap in science education research. In the next section, our theoretical conceptualization of what constitutes modeling competence will be briefly explained.

1.3 | Modeling competence in science education

In general, modeling competence is seen as one necessary element of teachers' professional competences (Günther et al., 2019; Osborne, 2014). Modeling competence is defined as a system of the knowledge, skills, and abilities necessary to engage in the process of developing and using models for reasoning in science, including motivational dispositions to apply these capabilities for problem-solving in specific situations (Upmeier zu Belzen, van Driel, et al., 2019). When achieving modeling competence, a person should better understand scientific concepts, develop an appreciation of the nature of science, and advance in their mastery of the scientific process (Gilbert & Justi, 2016).

As suggested in recent literature (Chiu & Lin, 2019; Nicolaou & Constantinou, 2014; Nielsen & Nielsen, 2021a), modeling competence can be divided into three dimensions: *meta-modeling knowledge*, *modeling practice*, and the *modeling product*. In the following, we will expand on each of these dimensions, highlighting the theoretical scope of this article.

1.3.1 | Meta-modeling knowledge

Meta-modeling knowledge is a term commonly used to describe epistemic ideas about models and modeling. Schwarz et al. (2009) defined meta-modeling knowledge as knowledge about "how models are used, why they are used, and what their strengths and limitations are"

TABLE 1 Theoretical framework for meta-modeling knowledge (Upmeier zu Belzen, van Driel, et al., 2019)

Aspect	Level I	Level II	Level III
Nature of models	Replication of the phenomenon	Idealized representation of the phenomenon	Theoretical reconstruction of the phenomenon
Multiple models	Different model objects	Different foci on the phenomenon	Different hypotheses about the phenomenon
Purpose of models	Describing the phenomenon	Explaining the phenomenon	Predicting something about the phenomenon
Testing models	Testing the model object	Compare the model and the phenomenon	Testing hypotheses about the phenomenon
Changing models	Correcting defects in the model object	Revising due to new insights	Revising due to the falsification of hypotheses about the phenomenon

Note: This framework uses the term *model object* referring to the work of Mahr (2011). In the present article, the term modeling product is used instead, because this term is more established in science education literature (e.g., Chiu & Lin, 2019).

(pp. 634–635). There are different approaches to conceptualize meta-modeling knowledge, each of them defining related aspects (Schwarz et al., 2009) or dimensions (Crawford & Cullin, 2005) as part of meta-modeling knowledge. One of the more common frameworks for meta-modeling knowledge, which will be used throughout this study, was proposed by Upmeier zu Belzen, van Driel, et al. (2019). They propose five aspects of meta-modeling knowledge: *nature of models*, *multiple models*, *purpose of models*, *testing models*, and *changing models*. For each, three levels of understanding are distinguished (Table 1). Level I is related to naïve views, understanding models as direct copies of reality and focusing on features of the modeling product itself, rather than on the representational and predictive function of models. Level II is related to more advanced views, understanding models mainly as idealized representations or media to visualize and explain something, while level III adds the appreciation of the predictive power of models as research tools (Upmeier zu Belzen, van Driel, et al., 2019).

Ke and Schwarz (2020, p. 5) distinguish between meta-modeling knowledge independent of the specific learning context (i.e., decontextualized meta-modeling knowledge) and "epistemological knowledge about models and modeling in action" (i.e., contextualized meta-modeling knowledge). The context-dependency of meta-modeling knowledge has been suggested to be an issue of critical importance for assessing and teaching meta-modeling knowledge, which, therefore, should be further investigated in science education (Krell et al., 2014; Sikorski, 2019).

1.3.2 | Modeling practice

Science education literature addresses modeling practice in several theoretical frameworks that use overlapping terms such as modeling activities (Fretz et al., 2002; Göhner & Krell, 2020a), model-based learning practices (Louca & Zacharia, 2012), or modeling phases (Constantinou, 1999). In this article, we will refer to *modeling practice* as an umbrella term encompassing any modeling behavior or cognitive operation while being engaged in modeling. To the specific modeling practices carried out by the participants in this study, however, we will refer to as *modeling processes*, which are

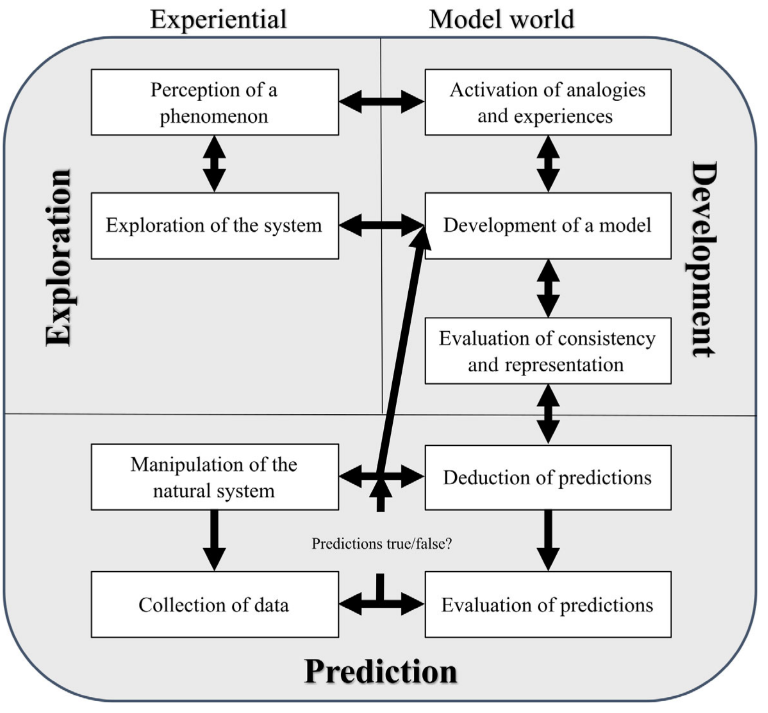

FIGURE 1 Model of modeling (adapted from Krell et al., 2019)

specific sequences of observable and distinguishable operations, in turn, referred to as *modeling activities* (Göhner & Krell, 2020a).

These modeling activities are incorporated in the model of modeling (Figure 1), which is the theoretical framework used for the description of modeling processes in this article (see Krell et al. (2019) for a detailed description).

The model of modeling distinguishes between the experiential world and the modeling world (Göhner & Krell, 2020a; Krell et al., 2019) and aligns with the model of scientific activity (NRC, 2012). Collecting data, making observations, and conducting experiments are part of the real world "investigating" sphere (i.e., experiential world), while models are part of the theoretical "developing explanations and solutions" sphere (i.e., modeling world). Additionally, the role of models as epistemic tools (Knuuttila, 2011) is emphasized in the model of modeling, when new hypotheses are deduced or predictions are drawn from the model, which are then also empirically tested (Dounas-Frazer & Lewandowski, 2018). Practically, the model of modeling can be used to operationalize the assessment and description of modeling processes.

1.3.3 | Modeling product

The main outcome of any modeling process is the development of a tangible, visible, and communicable artifact (i.e., modeling product) that demonstrates the modeler's understanding and that can be evaluated by specific criteria for its quality (e.g., epistemic criteria; Pluta et al., 2011). In learning contexts, modeling products are understood to be of high quality, if they externalize and express learners' thoughts and help them visualize and examine components of their theories (Jonassen & Ionas, 2008). Commonly in science education research, modeling products are evaluated in a more content-related approach based on the integration of specific components or relationships between these components (Chang et al., 2020). Chiu and Lin (2019) identified "a lack of deep discussions on the topic of modeling products" (p. 2). Most available studies analyze students' modeling products as indicators for evaluating students' modeling practices and meta-modeling knowledge (Bamberger & Davis, 2013; Cheng & Lin, 2015; Ergazaki et al., 2007; Schwarz et al., 2009).

Relevant to this study, modeling products can be evaluated for their quality, that is including the components and relationships that are required to accurately explain and predict the phenomenon, and evaluated for their complexity, that is the number of components and relationships in the model (Göhner & Krell, 2020a).

1.3.4 | Connecting meta-modeling knowledge, modeling practice, and modeling product

Many researchers in science education propose modeling practices and meta-modeling knowledge as the two broad constituent dimensions of modeling competence (e.g., Nicolaou & Constantinou, 2014; Upmeier zu Belzen, van Driel, et al., 2019). Chiu and Lin (2019) propose to add the modeling product as a third dimension as competent modelers are assumed to develop the ability to construct high quality models. Some researchers claim, that meta-modeling knowledge shapes or guides the practice of modeling (Lee & Kim, 2014; Nicolaou & Constantinou, 2014; Schwarz et al., 2009). Others conceptualize it the other way around, proposing that engagement in the modeling practice might contribute to the development of meta-modeling knowledge (Gobert & Pallant, 2004; Schwarz & White, 2005). Empirical findings propose that engagement in the modeling practices alone is not sufficient to foster meta-modeling knowledge but that associated reflections on the practices are necessary to reach this goal (Schwarz & White, 2005). This was also found in studies focusing on the broader construct of scientific meta-knowledge (i.e., nature of science; Abd-El-Khalick & Lederman, 2000). However, only a few studies had put the suggested relationships between the dimensions of modeling competence to the test (Chiu & Lin, 2019; Sins et al., 2009).

Summarizing, meta-modeling knowledge, and modeling products have been studied by many researchers and defined in a variety of approaches. However, process-oriented studies of students' or teachers' engagement in modeling, opposed to meta-modeling knowledge and modeling products, are still widely lacking in science education (Louca & Zacharia, 2012; Nicolaou & Constantinou, 2014). Therefore, we set to identify and characterize meta-modeling knowledge, modeling practices, and quality of modeling products in a sample of secondary preservice biology teachers and investigate the relationships between these three dimensions of the modeling competence. This will provide valuable theoretical knowledge, which is also important for understanding how to promote these dimensions in science teacher education (Nicolaou & Constantinou, 2014).

1.4 | Aims and research question

The following research question will be addressed in this study: What are the relationships between secondary preservice biology teachers' meta-modeling knowledge (contextualized and decontextualized), their modeling practices, and the quality of their modeling products while engaging in a modeling task? Following common assumptions in the literature, we expect all dimensions of modeling competence to be positively related.

2 | METHODS

2.1 | Context of the study

This study is situated in the first phase of secondary preservice biology teacher education in Germany. Secondary preservice biology teachers in Germany usually study two subjects (with one of them being biology) in a six-semester bachelor's program, followed by a four-semester master's program (concurrent teacher education programs). At the end of their studies, secondary preservice biology teachers are requested to having developed basic professional knowledge and competences needed for their profession (Neumann et al., 2017). These include knowledge and competences regarding inquiry and reasoning in science (KMK, 2019). A significant positive development of German secondary preservice science teachers' scientific reasoning competences over their course of studies has been described in empirical studies (Krüger et al., 2020).

2.2 | Sample

The sample population consisted of secondary preservice biology teachers, enrolled in the bachelor or master teacher education program in one of two involved German universities. To increase the likelihood of observing a variety of different modeling processes, a theoretical sampling strategy—heterogeneous sampling (Patton, 1990)—was used in the present study. Previous studies suggest that the quality of scientific modeling is positively related to scientific reasoning competences (e.g., Cheng & Lin, 2015) and general cognitive abilities (e.g., Nehring et al., 2015). Therefore, these two variables were used as screening variables, employing established pen-and-paper instruments (Krüger et al., 2020; Liepmann et al., 2007). Fifty-seven secondary preservice biology teachers, who had extreme scores (one half standard deviation higher or lower than the mean scores of the respective norm sample) in both assessments, were invited to participate in the study. Thirty-five secondary preservice biology teachers agreed to participate, aged between 17 and 39, with an average age of 24 years. Twenty-two participants were enrolled in the bachelor's program when participating in the study and 13 were enrolled in the master's program. Eight of the secondary preservice biology teachers were additionally enrolled in another scientific subject (including chemistry, physics, food science, agricultural science, and computer science). The study was not mandatory for any university courses or obligatory parts of the curriculum; participation was voluntary. Researchers and participants had no formal relationships to one another.

2.3 | Black box modeling task

A black box modeling task was applied in this study. The black box approach is established in science education research to study processes of scientific thinking and modeling (Lederman & Abd-El-Khalick, 2002; Passmore & Svoboda, 2012). In this study, a water black box was used

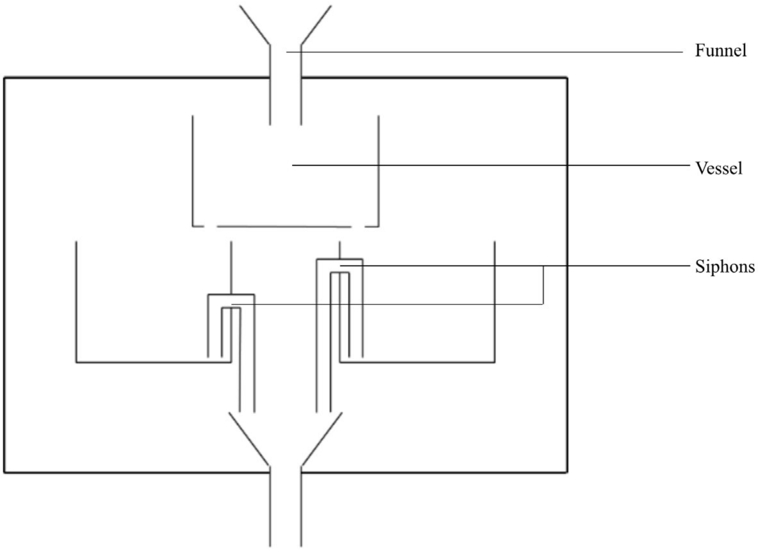

FIGURE 2 Model of the inner mechanism of the water black box used. Water is funneled into a first vessel and then directed equally into two further vessels. If a specific volume of water is reached in each of these vessels, the water flows out the black box through a siphon. As the siphons are installed on different heights, the specific volume of water for each vessel differs (Krell et al., 2019)

(Figure 2). It can be explored by filling the black box with water (input), which then results in measurable outputs of water (see Krell et al., 2019 for a detailed description of the black box). It has been shown that black box approaches are suitable to elicit modeling processes, in which models are used as epistemic tools to investigate the black box (Krell et al., 2019; Passmore & Svoboda, 2012), including all steps considered in the model of modeling (Figure 1).

Giving an example, repeating an input of 400 ml six times, produces a pattern of 0, 400, 600, 400, 0, 1000 ml outputs. Typically, the third output cannot readily be explained by the participants. In the model of modeling this resembles the perception of a phenomenon, which is a starting point for modeling processes as described above, including the exploration of the system (i.e., the black box), the activation of analogies and experiences (of what might be inside the black box), the development of a (drawn) model and its evaluation regarding consistency and representation. Using imagistic simulation, hypotheses (e.g., about the next output) can be deduced from the drawn model and the model can be tested by making another input, leading to the model's confirmation or rejection. In the latter case, the drawn model should be modified.

To gain additional insights into the participants' reasoning processes, they were asked to concurrently think aloud (Leighton & Gierl, 2007). This was practiced with three short exercises as part of the introduction to the study, in which the participants also answered the questionnaire on decontextualized meta-modeling knowledge (see below), were informed that participation was voluntary, and signed an informed consent form. After these preparations, the participants were brought into a room equipped with three video cameras, the black box, some prefilled beakers of water, a bucket as water reservoir, and a chalkboard. The first author briefly explained the basic functionality of the black box using a prepared script and provided the following task: "Draw a model of the inside of the black box." Participants were informed, that there are no time constraints. The first author stayed in the room to prevent any technical errors and, if necessary, to remind the participant to concurrently think aloud; otherwise, he did not intervene.

2.4 | Data collection and analysis

All qualitative data analyses were done within the methodological frame of qualitative content analysis, that is including category systems guiding analysis, coding by different persons, finding consensus coding after discussion, and calculating Cohen's Kappa (κ) as a measure of intrarater- and interrater-agreement (Schreier, 2012). In addition, the following procedures of data collection and analysis have been conducted in this study.

2.4.1 | Decontextualized meta-modeling knowledge

The participants' meta-modeling knowledge was assessed based on the framework described above (Upmeier zu Belzen, van Driel, et al., 2019), using an established pen-and-paper questionnaire which consists of five constructed response items (Krell & Krüger, 2016). These items are related to the five aspects *nature of models*, *multiple models*, *purpose of models*, *testing models*, and *changing models* (Table 2). The questionnaire was given to each participant immediately before engaging in the black box modeling task. In the questions, the respondents were asked to provide their understanding of the five aspects related to the scientific discipline of biology; however, as no more specific context was provided in the tasks and the questions have not been answered during any kind of modeling activity, we will refer to what has been assessed in the questionnaire as *decontextualized meta-modeling knowledge* (Ke & Schwarz, 2020). In total, 166 out of the possible 175 responses were analyzed from the 35 participants, as some participants did not respond to all five questions or responses could not be assigned to any level. All responses to the questionnaire were qualitatively analyzed using an already established category system, by which a specific level can be identified (ranging from I to III; Table 1) in the responses for each of the five aspects (Krell & Krüger, 2016). Statements including multiple response levels were always given the highest identified level. Each statement was coded twice by the first author and additionally by a trained student assistant, achieving substantial intrarater-agreements ($\kappa = 0.68$), and substantial interrater-agreements ($\kappa = 0.71$).

2.4.2 | Contextualized meta-modeling knowledge

The transcripts of the participants' modeling processes were analyzed by a trained student assistant based on the same category system that has been used for the analysis of the questionnaire (Table 2). Statements related to the five aspects of meta-modeling knowledge, verbalized by the participants throughout their modeling process, were identified and coded accordingly. Hence, this procedure aimed to assess the participants' *contextualized meta-modeling knowledge* or their meta-modeling knowledge "in action," respectively (Ke & Schwarz, 2020). In line with the evaluation of the questionnaire, participants expressing statements of varying levels throughout their modeling process were given the highest level observed for each aspect. To secure the quality of the analysis, 20% of the material (7 of the 35 transcripts) were randomly selected and

TABLE 2 The constructed response items for the five aspects of meta-modeling knowledge, with excerpts of response examples provided for each level and aspect

Item	Level I	Level II	Level III
Nature of models			
Explain the extent to which biological models are equivalent to their biological phenomenon.	A model should be as close to the original as possible and represent it well. (Daphne)	A model resembles the original strongly to little. Depending on the purpose of the model, it has structural or functional similarities. (James)	Models are just theoretical constructs created by people. (Selina)
Alternative models			
Why are there different models of one biological phenomenon?	As soon as a model is equivalent to the original, multiple perspectives are possible in model development and representation. (Kara)	As some models represent only a specific function or process, there are multiple possibilities to resemble an original object. (Celine)	There can be multiple models for one original object if different theories describing that model exist or if there are various hypotheses, which have neither been falsified nor verified yet. (Cynthia)
Purpose of models			
What is the purpose of models in biology?	Models in biology serve various purposes; in my opinion, however, they have a superficial purpose: to allow observations of the initial object that would not be possible under normal circumstances. (Claudia)	[Models] explain the structures of the source object. (May)	[The purpose of models is the] simulation of biological processes to gain knowledge or to solve problems. (Jonathan)
Testing models			
How do we test whether a biological model serves its purpose?	By testing the model, that is, using it in biology classes and testing it. (Frida)	By comparing with the original and see if the model shows/explains everything. (Achilles)	The model can be used to test hypotheses/make predictions that can be evaluated. (Ryan)

TABLE 2 (Continued)

Item	Level I	Level II	Level III
Changing models			
What are some reasons for changing a biological model?	Depending on its use, that is, if the model is understood in class or not, it is important to further adapt the model and improve it. [...] It must be a refinement or correction. (Frida)	There are new data that indicate errors in the previous model. (Carlo)	Predictions made by the model are not correct. (Floyd)

Note: Names in brackets are pseudonyms. Note that the questionnaire has been developed and administered in the German language. The full German questionnaire including the category system is available upon request from the third author.

coded twice over an interval of two weeks by the trained student assistant and once by the first author. Cohen's Kappa indicated substantial intrarater- ($\kappa = 0.69$) and substantial interrater-agreement ($\kappa = 0.72$).

2.4.3 | Modeling practice

For analyzing modeling practices, the behavior and verbalizations of the participants engaging in the black box modeling task were videotaped. The verbalizations were then transcribed verbatim, including selected behavioral aspects (e.g., making an input or observing an output). These transcripts were initially analyzed qualitatively to identify single modeling activities. For this, an established category system was used in the analysis. It consists of 9 main categories and 19 subcategories representing the modeling activities (Table 3). The nine main categories of the category system stem from the model of modeling described earlier in this study (Figure 1; for a detailed description see Krell et al., 2019). If necessary, the videos were considered as additional data sources in the analysis. Each transcript was coded twice over an interval of two weeks by the first author and additionally by a trained student assistant. Cohen's Kappa indicated almost perfect intrarater-agreement ($\kappa = 0.83$) and substantial interrater-agreement ($\kappa = 0.77$).

From the resulting sequences of modeling activities, state transition graphs were built to visualize each participant's modeling process (Andrienko & Andrienko, 2018). Each participant's state

TABLE 3 Activities of the modeling process (Krell et al., 2019)

Phase	Category	Sub-category (activity)
Exploration	1. Perception of a phenomenon	
	2. Exploration of the system	2.1. Input/output (exploratory)
		2.2. Summarizing/describing observations
		2.3. Input/output (pattern detection)
		2.4. Recognizing patterns
	3. Activation of analogies and experiences	
Development	4. Development of model	4.1. Graphically develop model
		4.2. Change model to optimize consistency
		4.3. Change model to optimize representation
		4.4. Reject model due to poor consistency/ representation
	5. Evaluation of consistency and representation	5.1. Evaluate consistency
		5.2. Evaluate representation
	6. Finding of consistency and representation	
Prediction	7. Deduction of predictions	
	8. Evaluation of predictions	8.1. Input/output (to test predictions)
		8.2. Confirmation of prediction
		8.3. Falsification of prediction
	9. Modification/rejection of model	9.1. Change model due to falsified predictions
		9.2. Reject model due to falsified predictions

transition graph consists of knots for each state (i.e., modeling activity) and edges depicting the transitions between two modeling activities. The edges are assigned the number of transitions observed between each modeling activity as weight.

All modeling processes were qualitatively evaluated based on the modeling activities shown during the process and their sequential order visualized as state transition graphs. Additionally, the state transition graphs were used to quantify the modeling practices for further analysis. Two variables, the complexity and homogeneity of the secondary preservice biology teachers' modeling processes, were estimated as described in Göhner and Krell (2020a, 2020b). This is shortly described in the following.

Modeling processes are understood to be more complex, if they include various modeling activities (Göhner & Krell, 2020a). Complexity was therefore estimated using the graph metric known as "communities," which involves counting all subgraphs within each state transition graph (Porter et al., 2009). Participants may show a limited range of activities, resulting in state transition graphs with more subgraphs, as some activities are not addressed and thus, not connected to others. The basic communities score was further normalized by subtracting the communities score of each participant from the maximum communities score achieved in this study.

Modeling processes are understood to be more homogenous, if the observed modeling activities and the transitions between them are more equally distributed, leading to the state transition graph being more independent from a single knot (Göhner & Krell, 2020a). To estimate homogeneity, the centrality of every state transition graph was determined. The centrality score was reciprocally transformed, as centrality describes the dependence of a graph on a single knot (Newman, 2010). To account for structural outliers, homogeneity was estimated using the sum of the three measures of centrality: closeness centrality, degree centrality, and betweenness centrality (Ronqui & Travieso, 2015).

2.4.4 | Modeling product

To assess the quality of the participants' modeling products, a novel category system was inductively developed, based on the previous work of the authors. As suggested by Bielik et al. (2018), modeling products were considered of high quality when they included the components and relationships that are required to accurately explain the black box phenomenon (Figure 2). Here, modeling products of high quality were expected to include three concepts, which were found to be necessary to explain the water distribution inside the black box and the input and output patterns (Krell et al., 2019): (1) The modeling product should contain water reservoirs filling up with water, (2) the water reservoirs should be embedded in a parallel system of diverging paths, and (3) water should be fully emptied at a specific fill level. The specific realization of these three concepts can vary, incorporating alternative ideas; water reservoirs, for example, could be realized through vessels, cisterns, or even sponges. As some modeling products could not be evaluated based on their drawn appearance alone, the participants' verbal statements during the development of their modeling products were considered as an additional data source.

The participants' final modeling products were photographed, digitally reconstructed, and coded. However, in two cases, participants presented multiple final modeling products as possible solution for the given black box task. To code the final modeling products, it was noted if each concept was integrated, by being either drawn or described verbally. Participants' scores range from zero to three, integrating neither of these concepts (=0), one (=1), two (=2), or all of these concepts (=3) into their drawn modeling product or verbal descriptions thereof. Further attributes of the modeling products like their aesthetic appearance or fit to the obtained data have not been evaluated. For participants with multiple final modeling products, only the highest scoring modeling product was considered in the statistical analysis. Every modeling product ($N = 42$) was coded twice over an interval of two weeks by the first and once by the second author. Cohen's Kappa indicated almost perfect intrarater- ($\kappa = 0.84$) and almost perfect interrater-agreement ($\kappa = 0.82$).

2.4.5 | Relationships between meta-modeling knowledge, modeling process, and modeling product

Spearman's rank correlations were calculated to analyze the relationship between decontextualized and contextualized meta-modeling knowledge for each aspect, complexity, and homogeneity of the modeling process, and the modeling product score. To obtain a measure of decontextualized and contextualized meta-modeling knowledge, the participants' mean score across the five aspects were calculated. Furthermore, due to the rather small sample size, nonparametric tests were carried out. Effect size measures have been calculated based on Lenhard and Lenhard (2016).

To provide in-depth insights into the potential nature and direction of the relationships between secondary preservice biology teachers' meta-modeling knowledge, modeling practices, and modeling products, sample cases will be qualitatively described. Hereby, the cases have been selected based on the statistical relationships found in the quantitative analysis. The selected cases were analyzed by reconstructing their individual modeling processes from the transcripts, which are illustrated as codelines, showing the sequential order of the individual modeling activities.

3 | RESULTS

Aggregated scores for meta-modeling knowledge, modeling processes, and modeling products will be provided in the following section. The full table with the individual scores can be found in Table S1.

3.1 | Decontextualized meta-modeling knowledge

Table 4 shows the distribution of response levels regarding the five aspects of meta-modeling knowledge assessed with the questionnaire. Across the five aspects, the mean response levels range from 1.97 for the aspect *nature of models* to 2.31 for the aspect *multiple models*. The mean

TABLE 4 Distribution of response levels regarding the fives aspects of decontextualized meta-modeling knowledge

Aspect	*N*	*n* responses level I	*n* responses level II	*n* responses level III	Mean response level (±*SD*)
Nature of models	34	4	27	3	1.97 (±0.45)
Multiple models	35	1	22	12	2.31 (±0.52)
Purpose of models	33	7	13	13	2.18 (±0.76)
Testing models	30	2	18	10	2.27 (±0.57)
Changing models	34	3	24	7	2.12 (±0.53)

TABLE 5 Distribution of response levels regarding the fives aspects of contextualized meta-modeling knowledge identified within the modeling processes

Aspect	*N*	*n* responses level I	*n* responses level II	*n* responses level III	Mean response level (±*SD*)
Nature of models	21	4	2	15	2.43 (±0.79)
Multiple models	7	3	0	4	2.14 (±1.07)
Purpose of models	16	2	10	4	2.13 (±0.62)
Testing models	15	0	5	10	2.67 (±0.48)
Changing models	11	2	4	5	2.27 (±0.79)

level assigned to the 166 responses from all participants is 2.17, suggesting, that the participants understand models mainly as idealized representations or media to visualize and explain something (Upmeier zu Belzen, van Driel, et al., 2019).

3.2 | Contextualized meta-modeling knowledge

One-hundred twenty-five statements related to meta-modeling knowledge were identified in the modeling processes of the 35 participants, which could be assigned to response levels of the category system (Table 5). It should be noticed that the number of statements related to meta-modeling knowledge in the modeling processes varies between the participants and the aspects mentioned, depending on the length, and complexity of the modeling processes. For some of the participants, up to sixteen statements related to meta-modeling knowledge were identified, whereas 19 participants rarely verbalized anything related to meta-modeling knowledge, leading to the identification of less than 3 statements for each of these participants. Six participants made no statements related to meta-modeling knowledge during their modeling processes at all. Also, the number of statements differs between the aspects of meta-modeling knowledge (Table S1).

Across the five aspects of meta-modeling knowledge, the mean response levels ranged from 2.13 for the aspect *purpose of models* to 2.67 for the aspect *testing models* (Table 5). The mean level of contextualized meta-modeling knowledge was 2.39, indicating a slightly higher understanding compared to the decontextualized meta-modeling knowledge. However, this still suggests that most participants understand models as idealized representations or media for visualization and explanation.

Figure 3 compares the mean response levels of decontextualized and contextualized meta-modeling knowledge for each aspect. A Wilcoxon signed-rank test revealed a significant difference between the participants' level of decontextualized and contextualized meta-modeling knowledge for the aspect *nature of models*, with decontextualized meta-modeling knowledge being significantly lower ($z = -2.84$, $p = 0.005$; $d = 1.09$, large effect size measure). For the other aspects, no significant differences were found (i.e., $p > 0.05$).

In summary, the meta-modeling knowledge of the participants, independent of being decontextualized or contextualized, indicates an understanding of models as idealized representations or media for visualization and explanation as the predominant perspective on models in our sample.

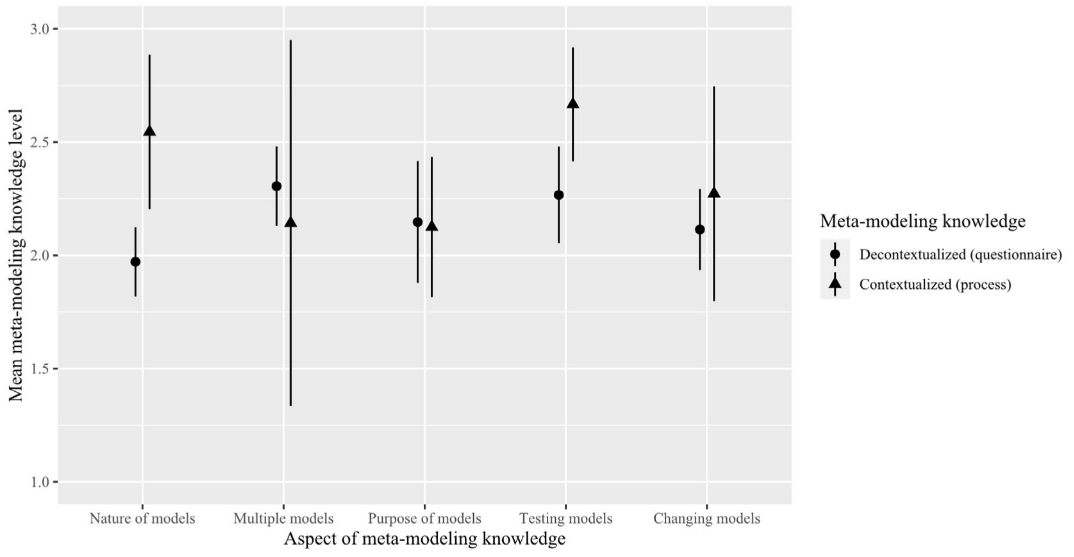

FIGURE 3 Comparison of mean meta-modeling knowledge levels across the five aspects of meta-modeling knowledge and with regard to the context of assessment, either decontextualized within the questionnaire or contextualized as identified within the modeling processes. Points and triangles indicate mean values and lines indicate the range of the mean value ± two times standard error

3.3 | Modeling practice

As described above, the specific modeling practices carried out by the participants in this study are referred to as modeling processes, which consist of different modeling activities. The length of the participants' modeling processes varied between eight minutes to almost 2 h (mean length: 1 h 9 min). In these modeling processes, the participants conducted between six and eighteen (mean: 12) different modeling activities. The qualitative analysis revealed, that the modeling processes of 20 participants included modeling activities of exploration and model development. The modeling processes of 14 participants additionally included activities of prediction. One participant's modeling process showed only modeling activities of exploration (i.e., no model development). By transforming these modeling processes into state transition graphs, the complexity (mean: 4.94) and homogeneity (mean: 1.50) of the modeling processes were quantified. Figure 4 illustrates three examples of state transition graphs. The state transition graph of Claudia's modeling process (Figure 4a) is characterized by rather low complexity, as she included only a limited range of modeling activities into her modeling process. In comparison to all transitions, the transitions between modeling activities 2.1 (input/output, exploratory) and 2.2 (summarizing/describing observations) occur more often, described by a medium homogeneity score. The state transition graph of James' modeling process (Figure 4b) is characterized by medium complexity, as he includes a higher number of different modeling activities in his modeling process. The transitions between his modeling activities are rather equally distributed, reducing the dependence of the state transition graph from specific knots (i.e., modeling activities), leading to a high homogeneity score. Finally, the state transition graph of *Raphael's* modeling process (Figure 4c) is characterized by high complexity, as he includes every modeling activity but one in his modeling process. However, his homogeneity score is medium, as he transitions between the modeling activities 2.2

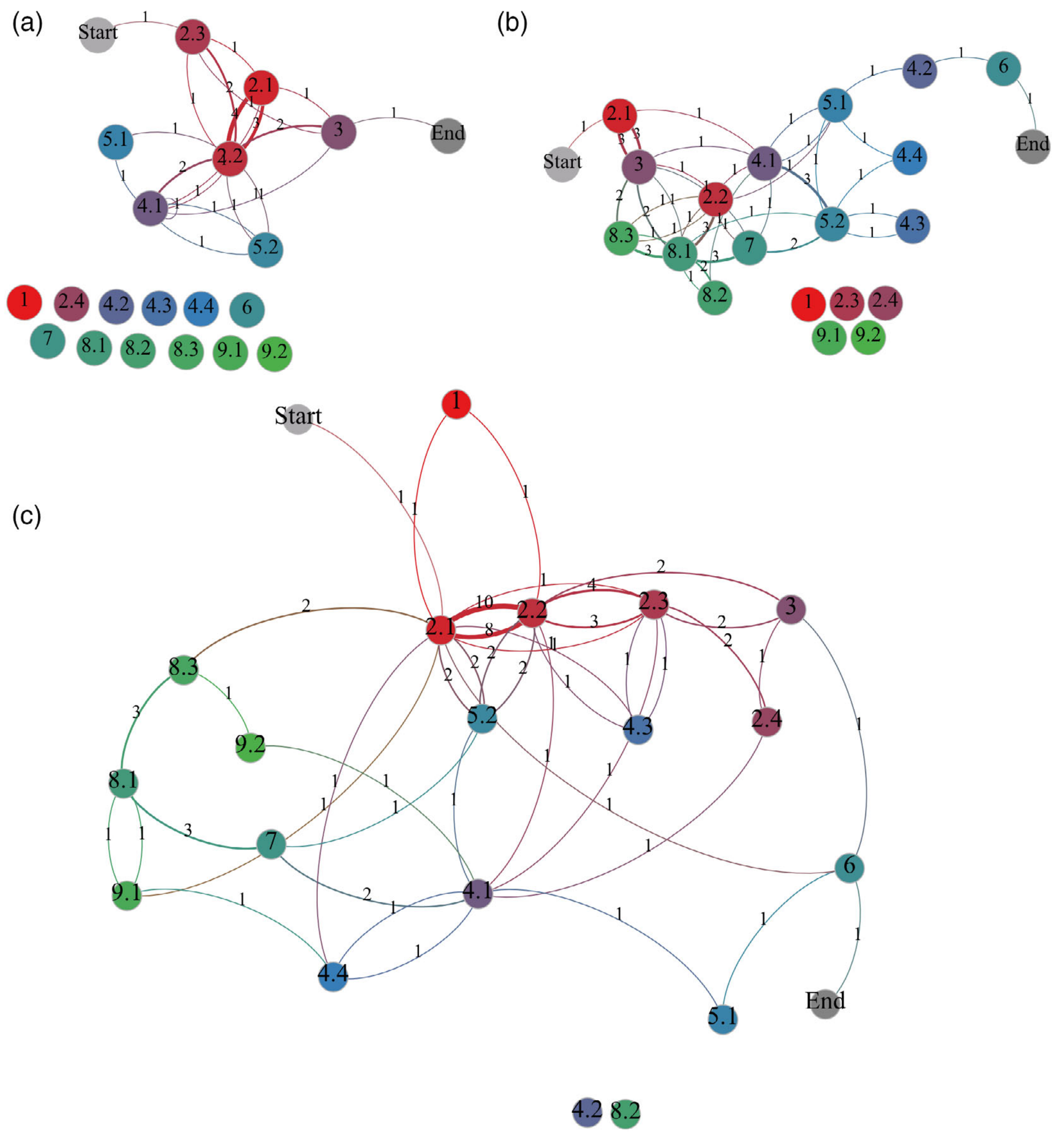

FIGURE 4 Three examples of state transition graphs for Claudia (a), James (b), and Raphael (c). The modeling activities (Table 3) are displayed as knots connected by their transitions, labeled with their occurrence

(summarizing/describing observations) and 2.3 (input/output, pattern detection) quite often in comparison to the other modeling activities. A complete state transition graph combining the modeling processes of all 35 participants can be found in the supplementary material (Figure S2).

A Mann-Whitney *U* test confirmed that participants with modeling processes, which included activities of prediction, reached significantly higher complexity scores ($M = 7.33$, $SD = 1.72$) than participants who did not include modeling activities of prediction ($M = 3.15$, $SD = 1.73$; $z = 4.70$, $p < 0.001$; $d = 2.62$, large effect size measure). Moreover, a correlation analysis (Spearman) between the complexity of the modeling processes and the length of the participants' modeling processes revealed a significant correlation ($r = 0.39$, $p = 0.021$; medium effect size measure). Other sample characteristics, like the subject combination of each participant or the enrollment status of the participants in either the bachelors' or masters' program, could not be shown to be systemically related to the modeling processes.

Summarizing, related to the modeling practices, the majority of our participants only engaged in modeling activities of exploration and model development, while activities of prediction were missing.

3.4 | Modeling product

Thirty-one participants produced a single modeling product at the end of their modeling processes. Two participants produced multiple models as possible solutions for the given task

TABLE 6 Integration of the three necessary concepts by the participants

	Concept 1 (water reservoirs filling up)	Concept 2 (parallel system of diverging paths)	Concept 3 (water reservoirs fully emptying at specific fill level)
Verbalized	32 (91.437%)	7 (20.00%)	11 (31.43%)
Drawn	32 (91.43%)	9 (25.71%)	3 (8.57%)
Verbalized and/or drawn	34 (97.142%)	9 (25.71%)	12 (34.29%)

TABLE 7 Overview of the number of concepts integrated by the participants verbally and/or drawn

	Participants (n) integrating the concepts		
Score (=number of concepts integrated)	Verbally	Drawn	Either verbally or drawn (=modeling product score)
0	2 (5.71%)	3 (8.57%)	1 (2.86%)
1	20 (57.14%)	21 (60.00%)	18 (51.43%)
2	9 (25.71%)	11 (31.43%)	12 (34.29%)
3	4 (11.43%)	0	4 (11.43%)

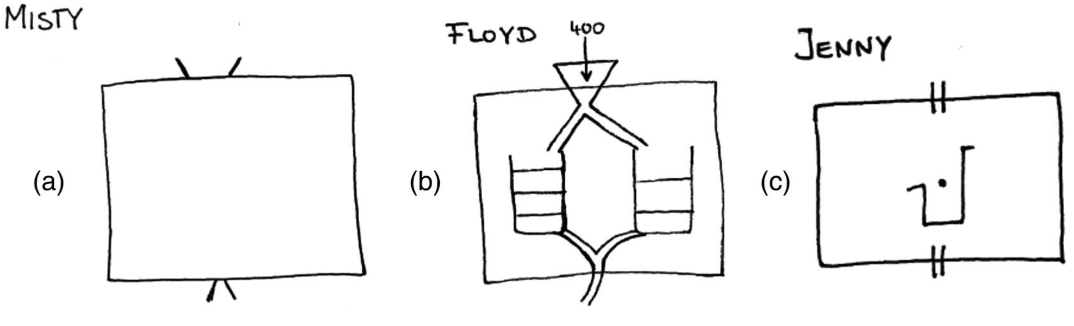

FIGURE 5 Examples of modeling products. Misty's modeling product was scored with 1 point, Floyd's and Jenny's modeling products with 2 points. Note that no participant drew a model that received full score (Table 6)

(*Claudia* presented five, *Kara* two). One participant (*Iris*) removed all her drawn modeling products before ending her modeling process, as she regarded none of the modeling products a suitable solution. Together with *Carlo*, she was one of two participants who had no drawn modeling product, although they verbalized their ideas.

The participants' modeling product scores ranged between zero and three with a mean score of 1.54 (SD = .73), indicating most participants were able to develop a model that integrates one or two target concepts. Only four participants reached full score. Table 6 shows how often each specific concept was integrated by the participants either verbally or drawn. Notably, no participant was able to *draw* all three target concepts (Table 7).

The first concept (water reservoirs filling up with water) was included in all models except one: *Misty* (Figure 5a) removed a vessel that she had drawn before from her final model, claiming "it is physically not possible to work like that," concluding that she cannot explain her data with only this one concept. Instead, she proposed—but did not draw—a mechanism at the bottom of the black box consisting of "something like a valve, regulated by pressure." Alternative concepts like these were found with nine other participants. However, these concepts were rarely evaluated as sufficient for the given task by the participants themselves. Of the remaining 33 models, seven also included the second concept (water reservoirs are embedded in a parallel system of two diverging paths). For example, *Floyd's* model (Figure 5b) showed two vessels on the bottom, filling up with water simultaneously. Though he did not draw the third concept (water reservoirs being fully emptied at a specific fill level), as he was unsure how a mechanism working like this could be drawn. Like *Floyd*, three other participants (*Alice*, *Raphael*, and *Susi*) drew a model that included the first two concepts but were only able to verbalize the third concept. In contrast, four participants (like *Jenny*; Figure 5c) were able to draw a mechanism, which included the third concept of water reservoirs being fully emptied at a specific fill level. This behavior was commonly explained with a tilting or turning mechanism. None of the participants drew a siphon, like it is actually used in the presented black box (Figure 2). Moreover, the second concept is drawn more often than explicitly verbalized, while the third concept is drawn rarely but more often verbalized.

Additionally, a Mann–Whitney U test regarding the participants' subject combinations revealed that secondary preservice biology teachers studying a second scientific subject developed modeling products with significantly higher modeling product scores ($M = 2.13$, $SD = 0.84$) as opposed to secondary preservice biology teachers with nonscientific secondary subjects ($M = 1.37$, $SD = 0.63$, $z = -2.34$, $p < 0.019$; $d = 0.86$, large effect size measure). Neither the length of the modeling processes, nor the participants' enrollment status in either the bachelors' or the masters' program could be shown to be systemically related to the quality of the modeling products.

In summary, participant's modeling products were of medium quality, mostly incorporating two of the three target concepts.

3.5 | Relationships between meta-modeling knowledge, modeling practice, and modeling product

The correlation analyses (Spearman) between the two measures of meta-modeling knowledge, the complexity and homogeneity of the modeling processes, and the modeling product scores revealed that there are no significant relationships between the five variables with one exception: the complexity of the modeling processes significantly correlates with the modeling product score (Table 8).

TABLE 8 Correlation table for Spearman's rank correlations between the variables decontextualized and contextualized meta-modeling knowledge, homogeneity, and complexity of the modeling processes, and the modeling product score

Variable	2	3	4	5
1. Decontextualized meta-modeling knowledge	0.27 (n.s.)	0.17 (n.s.)	0.21 (n.s.)	0.05 (n.s.)
2. Contextualized meta-modeling knowledge		0.25 (n.s.)	0.10 (n.s.)	−0.08 (n.s.)
3. Complexity			−0.01 (n.s.)	**0.41 (p <0.05)**
4. Homogeneity				0.11 (n.s.)
5. Modeling product score				

Note: $N = 29$ for correlation analyses including contextualized meta-modeling knowledge and $N = 35$ for the other analyses.

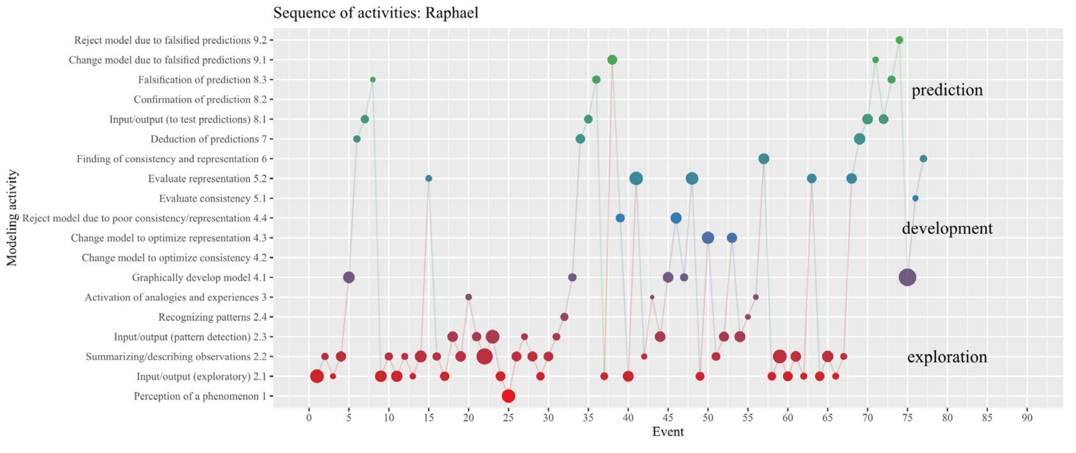

FIGURE 6 Sequence diagram showing the modeling process of Raphael in detail

Based on Mann–Whitney U tests, no significant differences between modeling processes which included modeling activities of prediction, and modeling processes which did not include modeling activities of prediction, were found for the variables decontextualized meta-modeling knowledge ($p = 0.61$), contextualized meta-modeling knowledge ($p = 0.19$), homogeneity ($p = 0.59$), and modeling product score ($p = 0.51$).

3.6 | Exemplary cases

In the following section, we present and discuss two exemplary cases to give more qualitative and holistic insights into the nature of the investigated relationships. With these two cases, we were especially interested in the only statistically significant relationship that was found between the modeling practices and the modeling products. Therefore, two cases with similar values along the complexity of the modeling processes and quality of the modeling products were chosen (case 1: *Raphael*, with high complexity score and high product score; case 2: *Angelina*, with medium complexity score and medium product score). These cases not only shed light on possible explanations for the found statistically significant relationship, but also generate insight into the other, statistically nonsignificant relationships as well.

3.6.1 | Example 1: Raphael

Raphael's modeling process (Figure 6) takes 1 h and 6 min and consists of 77 events. In his modeling process he shows a wide diversity of different modeling activities including exploration, model development, and prediction, and therefore received a high complexity score of 17 (Table S1).

He starts his modeling process by filling 400 ml of water into the black box twice, observing an output of roughly 400 ml after the second input. After documenting his observations on the board, he draws an initial modeling product (event 5) including the idea of an overflow vessel with a volume of 400 ml. From this modeling product, he deduces the hypothesis, that whatever his next input maybe, the output should be the same and tests his hypothesis by filling 100 ml of water into the black box. Contrary to his expectation, he cannot observe any output, which leads him to falsify and discard his modeling product as "obviously wrong." He then enters a longer phase (events 11–25) of modeling activities of exploration, collecting data by varying his input volumes and speeds until he recognizes a repeating pattern in the collected data (event 26). From this, he tries to develop a second modeling product retracing his collected data step by step while also simultaneously predicting the next output for each data point and testing it on the black box. Although this directly fails in the first step, as he predicts an output of 100 ml and observes only 50 ml, this leads him to the important observation, that the volume is somehow halved inside the black box. This motivates *Raphael* to modify his modeling product, adding a second compartment in which water is equally distributed. He then constantly switches between modeling activities of exploration and model development, collecting data, retrospectively explaining his collected data and modifying existing or developing new modeling products, which mostly add more compartments (events 35–50). Still not being able to explain all his observations, he concludes, that he needs a mechanism that "if a critical volume is reached, empties all the water or even more." Without integrating this idea into his drawn modeling product, he assumes that a very big input of 1300 ml would completely empty the black box, regardless of what is still left inside. Confirming this assumption, he verbalizes that the only thing unknown is the number and volume of compartments inside the black box (event 57). Therefore, with the black box now being empty again, he does a series of very small inputs comparing the outputs to his current modeling product (events 58–68). The now repeated observation of small inputs being halved, leads him to reason, that the black box contains two compartments. However, the small inputs did not help him in reasoning about the volume of each of the compartments, but he deduces the hypothesis "that the maximum volume of the black box should be around 1200 ml." After testing and falsifying this hypothesis (events 70–74), he claims "I am all out of ideas" and draws his final modeling product (Figure 7), which reaches the highest modeling product score. It includes the concepts

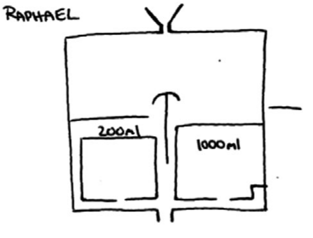

FIGURE 7 Modeling product of Raphael

1 (modeling product contains water reservoirs), 2 (water reservoirs are embedded in a parallel system), and 3 (water fully empties at a specific fill level), with concept 3 being verbalized, as he is unsure how to draw such a mechanism. After evaluating the appearance of his modeling product to an audience with "nobody will understand this, if I am not explaining it," he deems his modeling product sound "I think, the inside looks something like this." and ends his modeling process without further testing his modeling product to validate or falsify it.

Contrasting his modeling practice and modeling product, *Raphael's* answers to the questionnaire assessing decontextualized meta-modeling knowledge are all typical level 2 answers ($M = 2.00$), in which he consistently highlights models as idealized representations or media for visualization and explanation. During his modeling practice, only two statements showing his contextualized meta-modeling knowledge ($M = 2.50$) could be identified: In his first statement, he verbalizes a perspective on models more consistent with his answers to the questionnaire ("my model could be an explanation"; level 2). However, in his second statements he also makes clear, that he sees his model as a "hypothetical idea" (level 3), which better fits his modeling practice.

3.6.2 | Example 2: Angelina

Angelina's modeling process (Figure 8) takes 1 h and 1 min and consists of 62 events. In her modeling process, she shows a lower diversity of different modeling activities and only includes activities of exploration and model development. This leads to a medium complexity score (11).

Angelina starts her modeling process with one input of 400 ml. After observing no output, she inputs another 400 ml very slowly, resulting in the siphon inside the black box not emptying the vessel fully, but letting the 200 ml of water trickle out slowly. After documenting her observations, she follows up with a third, now faster, input, leading her to observe an output of 1000 ml as both vessels inside empty through their respective siphon. From this, *Angelina* verbalizes the idea of a vessel connected to a tilting mechanism (events 9–11). However, she does not attempt to draw a modeling product yet. Instead, she focuses on generating more data, following up with numerous, varying inputs, and observation of the outputs. She documents everything in a very thorough and systematic manner (events 12–26) until she claims to

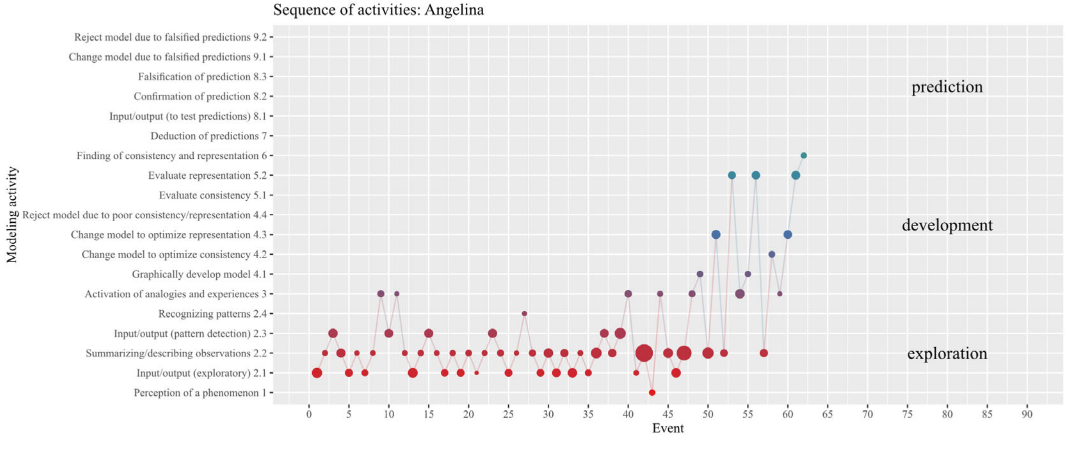

FIGURE 8 Sequence diagram showing the modeling process of Angelina in detail

recognize a reoccurring pattern. Still, she is unable to predict the next outputs from the previously collected data and continues to collect even more data, going into increasingly longer activities of summarizing her previous observations. After verbalizing the idea, that “to explain the varying outputs, multiple vessels with different volumes are necessary” she finally draws a modeling product, depicting three tilting vessels with different volumes. Reflecting on her “mathematically-oriented approach” and still not being able to explain her data with her modeling product, *Angelina* discards it. Activating the analogy of a wheel, she draws her final modeling product (Figure 9), consisting of a round, revolving vessel with different compartments. It reaches a medium modeling product score as it includes the concepts 1 (modeling product contains water reservoirs) and 3 (water fully empties at a specific fill level). With this modeling product she is now able to explain her inconsistent data. However, she expresses, that she is not convinced the modeling product is fully sound and can be seen as “provisional.” With this *Angelina* ends her modeling process.

Although, her scores regarding the modeling practice and modeling product are quite different than *Raphael's*, the scores of *Angelina* for decontextualized and contextualized meta-modeling knowledge perfectly match the scores of *Raphael*. Her answers to the questionnaire assessing decontextualized meta-modeling knowledge are also typical level 2 answers throughout ($M = 2.00$), like *Raphael*, highlighting models as idealized representations or media for visualization and explanation. Even her two statements assigned to contextualized meta-modeling knowledge ($M = 2.50$) are quite similar to *Raphael*: In her first statement, she is consistent with her answers in the questionnaire (“There are different containers, tilting and having different volumes, which explains the unregular volumes coming out”; level 2). In another statement, she reflects about her practice on a meta level, claiming that seeing her idea as a provisional model helps her thinking about it (level 3).

3.6.3 | Summary of the two cases

The cases of *Raphael* and *Angelina* illustrate two approaches to the modeling task at hand, which differ in their integration of the modeling product into the modeling practice: *Raphael* uses his modeling product as an epistemic tool by constantly switching between modeling phases of exploration, model development and predictions, in which he strongly relies on his modeling product to test his ideas. This test-driven modeling practice leads him to expose flaws in his modeling product, which he overcomes by then integrating new ideas (i.e., analogies) into his modeling product. Interestingly, while he shows this test-driven modeling practice with his initial, most simplistic modeling product of an overflow vessel (lowest modeling product score),

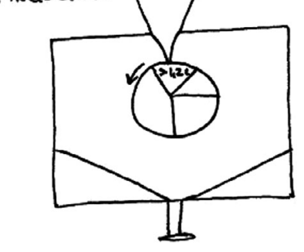

FIGURE 9 Modeling product of Angelina

he does not show this with his final modeling product (highest modeling product score). In contrast, *Angelina* approaches the modeling task using a more data-driven approach. After a long phase of exploration, collecting a lot of observations in a data table, she develops her first modeling product in the last third of her modeling practice attempting to explain the observed data retrospectively. While this also helps her in generating new ideas about the inner mechanism of the black box, she does not test her modeling product by drawing hypotheses and predicting outputs from the modeling product. Instead, she compares newly collected data to her previous observations and changes her modeling product to fit the data retroactively. Lastly, regarding their meta-modeling knowledge, both *Raphael* and *Angelina* achieve exactly the same scores for contextualized and decontextualized meta-modeling knowledge, although their modeling practice and product widely differ.

4 | DISCUSSION

In the present study, we set out to identify, characterize, and test the correlation between the three dimensions of modeling competence: meta-modeling knowledge, modeling practices, and modeling product. These three dimensions were operationalized as five variables, including decontextualized and contextualized meta-modeling knowledge for the meta-modeling knowledge dimension, complexity and homogeneity of the modeling processes for the modeling practices dimension, and a modeling product quality score. This approach aimed to uncover the relationships between the three dimensions, which literature proposes to be highly and positively related (Gobert & Pallant, 2004; Lee & Kim, 2014; Schwarz & White, 2005) but has not been investigated comprehensively yet (Chiu & Lin, 2019; Louca & Zacharia, 2012; Nicolaou & Constantinou, 2014).

4.1 | Meta-modeling knowledge

In line with earlier studies, the 35 secondary preservice biology teachers' decontextualized meta-modeling knowledge seems to be rather limited (Table 4), as their views of models in their answers to the questionnaire suggest mainly medial perspectives on models (Krell & Krüger, 2016; Torres & Vasconcelos, 2015). More sophisticated views of models, in which models are seen as epistemic tools that are being tested by deducing and testing predictions, are rarely addressed (i.e., level III; Krell & Krüger, 2016). However, theoretically less informed views about scientific models, reflecting a naïve understanding of models as copies of reality, were also only rarely observed, contradicting previous reports of these being quite common (Torres & Vasconcelos, 2015). For contextualized meta-modeling knowledge, it is suggested that specific contexts, including the black box modeling task used here, provoke more elaborate views of models (Ke & Schwarz, 2020). While this could be confirmed for the aspect *nature of models* (Figure 3), the average level of contextualized meta-modeling knowledge is only slightly higher compared to the decontextualized assessment and not significantly different statistically (Table 5). This may suggest that the specific black box modeling task used in this study emphasizes the hypothetical character of models, as the secondary preservice biology teachers are constructing their own hypothetical model, being made aware that models do not have to be final solutions or ready-made explanations (e.g., *Raphael*). Generally, the participants' decontextualized and contextualized meta-modeling knowledge varies between the different aspects of meta-modeling knowledge, suggesting unstable views about scientific models (cf., Krell et al., 2014).

4.2 | Modeling practice

Regarding the participants' modeling practices observed in this study, only some individual modeling processes can be seen as meaningful engagement, characterized by reflective insight, elaborate relational reasoning and monitoring the models' structure, and modeler's goals leading to the systematical testing of hypotheses and scoping of variables (Sins et al., 2005). While the correlation analysis between the complexity of the modeling processes and the length of the participants' modeling processes revealed a significant positive correlation, the nature of this relation does not become particularly clear, considering the qualitative data collected. However, it is obvious that with longer time on task, there is a higher probability of addressing more modeling activities. Still, only half of the modeling processes observed included activities of prediction. This is in line with previous studies suggesting that the predictive use of models is challenging for most sample groups, including teachers and students (Krell & Krüger, 2016; Passmore et al., 2014). Additionally, it is suggested in the literature that an evaluation of the modeling products by attempts of falsification or by systematically testing alternative hypotheses is also important for scientific modeling (Louca & Zacharia, 2015). Neither of these behaviors could be observed. A more in-depth discussion of our findings regarding the modeling practices including a comparison to other scientific practices like experimentation can be found in Göhner and Krell (2020a).

4.3 | Modeling product

In our sample group and for the given modeling task of the black box, the quality of the modeling products was rather low, with not a single participant drawing a modeling product incorporating all three necessary concepts (Table 6). This indicates a high difficulty of the given black box modeling task, which was initially chosen over an authentic and content-rich biological problem to reduce the influence of prior knowledge on the secondary preservice biology teachers' modeling processes. This advantage of the black box approach, on the other hand, is potentially limiting the generalizability of our findings and highlights the importance of context and content knowledge for modeling (Ruppert et al., 2017). This is further emphasized by some of our sample characteristics: secondary preservice biology teachers studying a second scientific subject developed modeling products with significantly higher modeling product scores than secondary preservice biology teachers with a nonscientific secondary subject. However, it is not evident from the qualitative data, if this correlation is grounded in the higher content knowledge, especially regarding physics knowledge, of participants studying two scientific subjects or if they are simply more familiar in working with scientific models. Regarding the latter, we rarely observed participants labeling and keying all the elements of their modeling products or developing explicit comparative modeling products, which was also observed by Bamberger and Davis's (2013) regarding student's modeling products prior to an instructional intervention.

4.4 | Relationships between meta-modeling knowledge, modeling practice, and modeling product

Investigating the relationships between the three dimensions reveals that not all relationships proposed in the literature were observed in the present study. While the relationship between modeling practices and modeling products was found to be statistically significant, no statistical evidence was found for the most commonly proposed relationship between meta-modeling knowledge and modeling products.

4.4.1 | The relationship between modeling practice and modeling product

In this study, a positive significant correlation between complexity of modeling processes and quality of modeling products was found (Table 8). This finding suggests a positive relationship between modeling practices and modeling products. This is reassuring, given that the analysis of modeling products is commonly used as a proxy for the more time-consuming analysis of modeling practice (Bamberger & Davis, 2013; Cheng & Lin, 2015; Ergazaki et al., 2007; Schwarz et al., 2009). However, the correlation coefficient indicates a medium effect size, suggesting a shared variance of about 17%. Most participants developed modeling products of rather low quality (Table 6 and Table S1). This matches their modeling processes of rather low complexity, which often did not include modeling activities of prediction. While in theory, this could be explained by low-quality modeling products limiting modelers in their modeling activities, making them unable to deduce suitable hypotheses and test them, no direct evidence for this causal relationship could be found in the qualitative data. Quite the contrary, the case of *Raphael* illustrates that even the most simplistic modeling product of a single overflow vessel can be used to deduce a hypothesis leading to the modeling product being tested and revised. Moreover, other observations suggest that the relationship between modeling practice and modeling product might be more complex. *Raphael*, ending his modeling process without testing his final highest-scoring modeling product in the same predictive manner, suggests that a high-quality modeling product does not automatically lead to complex modeling practices, although attempting a falsification of a developed model is considered to be an utterly important modeling practice (Giere et al., 2006).

Comparing *Raphael's* approach to the modeling task at hand with *Angelina's* approach, indicates that the *perceived soundness*—how we suggest to call it—of a modeling product may explain if and when a modeling product is used for predictions during the modeling practice. *Raphael* deems his modeling products as sound enough to reason with them and to test ideas, until his final modeling product is so sound it can be presented as a solution to the task. In contrast, *Angelina*, who shows a more data-driven approach and does not predict from her model at all, as she perceives all her modeling products, including the final one as not very sound (or "provisional"). She makes clear, that she is unsure in multiple instances, as she is not able to reproduce or to explain the black box phenomenon retrospectively until her final modeling product. This factor of perceived soundness could also explain, why *Angelina* does not draw a model at all during the first two thirds of her modeling practice, as she perceives her data as rather unsound until she recognizes a pattern, from which she is able to develop her first modeling product.

Furthermore, it is especially interesting that evidence for a relationship between modeling practices and modeling products was found by evaluating the modeling products based on the concepts they integrated (i.e., their content) and not by using epistemic criteria, like for example parsimony or conceptual coherence (Pluta et al., 2011). This adds more evidence to a potential influence of context on performance (Krell et al., 2014; Schwarz, 2002; Sikorski, 2019). Content knowledge (and therefore context) can be seen as a limiting or moderating factor, whose important role for scientific modeling has already been emphasized by different authors (e.g., Ruppert et al., 2017). As most of the participants in this study were studying biology as their only scientific discipline ($n = 30$), they might lack relevant physics knowledge to develop an appropriate modeling product for the inner mechanism of the black box. The task of discovering a black box is rather abstract, and hence, not representative for problem solving in authentic, content-rich scientific contexts (Leden et al., 2020). Therefore, it remains unclear how exactly domain-specific and situated content knowledge moderates the quality of the modeling products, as well as meta-modeling knowledge and modeling processes. Our findings emphasize a strong need for a systematic analysis of the role of content knowledge for modeling in research and the importance of context for interventions in educational settings.

4.4.2 | The relationship between meta-modeling knowledge and modeling practice

In general, statistical evidence for a relationship between meta-modeling knowledge and modeling practices, as well as meta-modeling knowledge and the two other variables considered in this study (Table 8), were surprisingly absent. This contradicts common assumptions of meta-modeling knowledge guiding the practice (Louca & Zacharia, 2012; Schwarz et al., 2009) and indicates that meta-modeling knowledge might not be a valid predictor for the quality of engagement in modeling practices and the modeling product. Of course, the small sample size of this study and the limitations discussed below have to be considered in the generalizability of these findings, but they clearly showthat more research regarding this relationship is needed.

Based on the average scores of meta-modeling knowledge, the present findings suggest that the secondary preservice biology teachers in our sample with rather low levels of meta-modeling knowledge are somehow able to engage in elaborate modeling practices. These modeling practices could be observed as modeling processes, which are characterized by high complexity and homogeneity scores, and include activities of prediction, using their models for reasoning about the black box (e.g., *Raphael*). On the contrary, some participants with high levels of meta-modeling knowledge, describing models consistently as hypothetical entities for knowledge generation, showed less elaborate modeling practices, including no activities of prediction in the herein observed modeling processes. These participants typically used their modeling products just to illustrate and communicate their solution (*Carlo*, see Table S1). Although Ke and Schwarz (2020) propose that meta-modeling knowledge in action might be stronger related with the practices (in contrast to abstract, decontextualized meta-modeling knowledge), the average contextualized meta-modeling knowledge was also not found to be significantly related to the complexity of the modeling processes. However, the response levels for both decontextualized and contextualized meta-modeling knowledge widely varied across the statements of single participants (e.g., *Claudia*, see Table S1), suggesting unstable views of models. In line with the observations regarding modeling products in this study, it seems likely that meta-modeling knowledge is activated by the participants based on the context of assessment and the aspect of meta-modeling knowledge they address. This is in concurrence to previous studies, discussing the context-dependency of meta-modeling knowledge and the construct of modeling competence as a whole (Gobert & Pallant, 2004; Krell et al., 2014; Sikorski, 2019; Sins et al., 2009).

The qualitative analyses of the presented cases of *Raphael* and *Angelina* shed some light on the relationship between meta-modeling knowledge and modeling practices. On the one hand, their direct comparison provides more evidence to the absence of a relationship between meta-modeling knowledge and modeling practices, as both show the exact same scores regarding meta-modeling knowledge but engage in completely different modeling practices. On the other hand, the case of *Angelina* clearly shows someone with consistent scores over all assessed dimensions of modeling competence, implying there might be a relationship, which just could not be found to be of statistical significance in this study. Qualitative analyses of the other cases, not presented in this study, indicate, that single responses regarding the aspects *testing models* and *changing models* often are in line with the observed modeling processes. However, these statements leave unclear whether meta-modeling knowledge guides the practice or engagement in practices promotes understanding (Gobert & Pallant, 2004; Schwarz et al., 2009). Moreover, our findings suggest that helping secondary preservice biology teachers improve their meta-modeling knowledge, for example in university courses, may not necessarily improve their modeling practice (Shi et al., 2021).

4.5 | Limitations of the study

This study has some limitations. First and foremost, the sample size is modest and only consists of secondary preservice biology teachers, which potentially limits the generalizability of our findings to this specific sample group. We hope our study can give an example to investigate other sample groups of science teachers. Further limitations must be considered regarding our methodology. As in most similar studies, meta-modeling knowledge (decontextualized and contextualized) was assessed as views on models (Krell & Krüger, 2016; Schwarz et al., 2009). Nicolaou and Constantinou (2014), though, distinguish between two dimensions of the overarching construct of modeling metaknowledge: the herein investigated meta-modeling knowledge (the epistemological awareness about the purpose and use of models) and metacognitive knowledge of the modeling process (the understanding of the actual modeling process).

Recently it has been suggested to also assess the latter (Lazenby et al., 2020). Yet to date, there is no established form of assessment in science education research, which addresses this dimension of metacognitive knowledge of the modeling process. The development of such an assessment tool based on a coherent and sound theoretical framework and independent of the traditional approaches to meta-modeling knowledge might yield further insights into the construct of modeling competence and its development (Nicolaou & Constantinou, 2014). Additionally, the participants in this study were not explicitly asked for verbalizing contextualized meta-modeling knowledge, leading to incomplete data sets, as not every participant verbalized thoughts about every aspect of meta-modeling knowledge throughout their modeling processes. 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). Furthermore, we cannot be sure if all reasoning processes of the participants are fully revealed by them concurrently thinking-aloud. For the quality of the modeling products, the participants were limited to develop their modeling products by drawing. Although this approach is quite common in science education research (Chang et al., 2020), modeling products could also be developed via computational modeling or hands-on modeling. The given black box modeling task, which allowed only for drawing, may have limited some of the participants in their modeling processes. Also, it should be acknowledged, that an evaluation of the modeling products based on epistemic criteria may yield further insights (Pluta et al., 2011). However, this was not applicable in the present study given the overall lack of complexity and variance of the modeling products.

Finally, this study solely focused on the modeling competence of secondary preservice biology teachers, without taking their ability to teach modeling competence to students into consideration, which is suggested to be even more challenging for teachers (Shi et al., 2021) and develops over vast time frames (Vo et al., 2019).

5 | CONCLUSION

The present study shows that the established assumptions, held by many science education researchers, about the dimensions that constitute modeling competence and their relationships, might not be empirically valid, at least for secondary preservice biology teachers. Firstly, our findings highlight that secondary preservice biology teachers' meta-modeling knowledge is not a reliable indicator for their engagement in modeling practices or for modeling competence in general. Secondly, the quality of modeling products seems to be a more reliable proxy assessment, as it corresponds to the observed modeling practice in this study. However, it becomes clear, that a valid assessment of secondary preservice biology teachers' modeling competence needs to consider meta-modeling knowledge, modeling practices, and modeling products holistically. Although more research is needed regarding the metacognitive knowledge of the modeling process or the influence of context on modeling, we hope that our study provides valuable insights for researchers, who work on an assessment of modeling competence. Untangling the construct of modeling competence further, could lead to further development of educational interventions, aiming to foster modeling competence in teachers and, consequently, improve the integration of models and modeling in classrooms worldwide.

ACKNOWLEDGMENTS

We thank the German Research Foundation (DFG) for funding the present project (award nr. 327507949). Any opinions, findings, conclusions, or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the DFG. Open Access funding enabled and organized by Projekt DEAL.

ORCID

Maximilian Felix Göhner https://orcid.org/0000-0003-4938-372X
Tom Bielik https://orcid.org/0000-0002-8227-5870
Moritz Krell https://orcid.org/0000-0003-2226-0383

REFERENCES

Abd-El-Khalick, F., & Lederman, N. G. (2000). Improving science teachers' conceptions of nature of science: Acritical review of the literature. *International Journal of Science Education*, *22*(7), 665–701.

ACARA (2015). *The Australian curriculum: science*. Retrieved from http://www.australiancurriculum.edu.au/

Andrienko, N., & Andrienko, G. (2018). State transition graphs for semantic analysis of movement behaviours. *Information Visualization*, *17*(1), 41–65.

Bamberger, Y. M., & Davis, E. A. (2013). Middle-school science students' scientific modelling performances across content areas and within a learning progression. *International Journal of Science Education*, *35*(2), 213–238.

Baumert, J., & Kunter, M. (2013). The COACTIV model of teachers' professional competence. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project* (pp. 25–48). Springer US.

BCMOE (2019). *Curriculum redesign*. Retrieved from https://curriculum.gov.bc.ca/rethinking-curriculum

Bielik, T., Opitz, S., & Novak, A. (2018). Supporting students in building and using models: Development on the quality and complexity dimensions. *Education Sciences*, *8*(3), 149.

Campbell, T., Oh, P. S., Maughn, M., Kiriazis, N., & Zuwallack, R. (2015). A review of modeling pedagogies: Pedagogical functions, discursive acts, and technology in modeling instruction. *Eurasia Journal of Mathematics, Science and Technology Education*, *11*(1), 159–176.

Carlson, J., & Daehler, K. (2019). The refined consensus model of pedagogical content knowledge in science education. In A. Hume, R. Cooper, & A. Borowski (Eds.), *Repositioning pedagogical content knowledge in teachers' knowledge for teaching science* (pp. 77–92). Springer.

Chang, H.-Y., Lin, T.-J., Lee, M.-H., Lee, S. W.-Y., Lin, T.-C., Tan, A.-L., & Tsai, C.-C. (2020). A systematic review of trends and findings in research employing drawing assessment in science education. *Studies in Science Education*, *56*(1), 77–110.

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.

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*, 307–323. https://doi.org/10.1007/s11165-019-09880-2

Chiu, M.-H., & Lin, J.-W. (2019). Modelling competence in science education. *Disciplinary and Interdisciplinary Science Education Research*. *1*, 1–11. https://doi.org/10.1186/s43031-019-0012-y

Constantinou, C. P. (1999). The Cocoa Microworld as an environment for developing modelling skills in physical science. *International Journal of Continuing Engineering Education and Life Long Learning*, *9*(3-4), 201–209.

Crawford, B. A., & Cullin, M. J. (2005). Dynamic assessments of preservice teachers' knowledge of models and modelling. In K. Boersma, M. Goedhart, O. de Jong, & H. Eijkelhof (Eds.), *Research and the quality of science education* (pp. 309–323). Springer.

Dounas-Frazer, D. R., & Lewandowski, H. J. (2018). The modelling framework for experimental physics: Description, development, and applications. *European Journal of Physics*, *39*(6), 64005.

Ergazaki, M., Zogza, V., & Komis, V. (2007). Analysing students' shared activity while modeling a biological process in a computer-supported educational environment. *Journal of Computer Assisted Learning*, *23*(2), 158–168.

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.

Giere, R. N. (1999). *Science without laws*. The University of Chicago Press.

Giere, R. N., Bickle, J., & Mauldin, R. (2006). *Understanding scientific reasoning*. Thomson Wadsworth.

Gilbert, J. K., & Justi, R. (2016). *Modelling-based teaching in science education*. Springer.

Gobert, J. D., & Pallant, A. (2004). Fostering students' epistemologies of models via authentic model-based tasks. *Journal of Science Education and Technology*, *13*(1), 7–22.

Godfrey-Smith, P. (2006). The strategy of model-based science. *Biology and Philosophy*, *21*(5), 725–740.

Göhner, M., & Krell, M. (2020a). Preservice science teachers' strategies in scientific reasoning: The case of modeling. *Research in Science Education*. https://doi.org/10.1007/s11165-020-09945-7

Göhner, M., & Krell, M. (2020b). Analyzing the relationships between pre-service biology teachers' modeling processes, scientific reasoning competencies and general cognitive abilities. In M. Ergazaki & K. Kampourakis (Eds.), *Contributions from biology education research*. Springer Nature.

Günther, S. L., Fleige, J., Upmeier zu Belzen, A., & Krüger, D. (2019). Using the case method to foster preservice biology teachers' content knowledge and pedagogical content knowledge related to models and modeling. *Journal of Science Teacher Education*, *30*(4), 321–343.

Harlow, D. B., Bianchini, J. A., Swanson, L. H., & Dwyer, H. A. (2013). Potential teachers' appropriate and inappropriate application of pedagogical resources in a model-based physics course: A "knowledge in pieces" perspective on teacher learning. *Journal of Research in Science Teaching*, *50*(9), 1098–1126.

Harré, R. (1970). *The principles of scientific thinking*. University of Chicago Press.

Jonassen, D. H., & Ionas, I. G. (2008). Designing effective supports for causal reasoning. *Educational Technology Research and Development*, *56*(3), 287–308.

Justi, R., & Van Driel, J. H. (2006). The use of the interconnected model of teacher professional growth for understanding the development of science teachers' knowledge on models and modelling. *Teaching and Teacher Education*, *22*, 437–450.

Ke, L., & Schwarz, C. V. (2020). Supporting students' meaningful engagement in scientific modeling through epistemological messages: A case study of contrasting teaching approaches. *Journal of Research in Science Teaching.*, *58*(3), 335–365.

Khan, S. (2011). What's missing in model-based teaching. *Journal of Science Teacher Education*, *22*(6), 535–560.

KMK (2019). *Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung [Common federal requirements for science and science education in teacher education]*. Retrieved from https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2008/2008_10_16-Fachprofile-Lehrerbildung.pdf

KMK (2020). *Bildungsstandards im Fach Biologie für die Allgemeine Hochschulreife [Educational standards in biology for the higher education entrance qualification]*. Retrieved from https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2020/2020_06_18-BildungsstandardsAHR_Biologie.pdf

Knuuttila, T. (2011). Modelling and representing: An artefactual approach to model-based representation. *Studies in History and Philosophy of Science Part A*, *42*(2), 262–271.

Krell, M., & Krüger, D. (2016). Testing models: A key aspect to promote teaching activities related to models and modelling in biology lessons? *Journal of Biological Education*, *50*(2), 160–173.

Krell, M., Upmeier zu Belzen, A, & Krüger, D. (2014). Context-specificities in students' understanding of models and modelling. In C. P. Constantinou, N. Papadouris, & A. Hadjigeorgiou (Eds.), *E-Book Proceedings of the ESERA 2013 Conference: Science Education Research For Evidence-based Teaching and Coherence in Learning*, Nicosia, Cyprus. pp. 1024–1035.

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.

Krüger, D., Hartmann, S., Nordmeier, V., & Upmeier zu Belzen, A. (2020). Measuring scientific reasoning competencies: Multiple aspects of validity. In O. Zlatkin-Troitschanskaia, H. A. Pant, M. Toepper, & C. Lautenbach (Eds.), *Student Learning in German Higher Education* (pp. 261–280). Springer.

Lazenby, K., Stricker, A., Brandriet, A., Rupp, C., Mauger-Sonnek, K., & Becker, N. (2020). Mapping undergraduate chemistry students' epistemic ideas about models and modeling. *Journal of Research in Science Teaching*. *57*, 794–824. https://doi.org/10.1002/tea.21614

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.

Lederman, N. G., & Abd-El-Khalick, F. (2002). Avoiding de-natured science: Activities that promote understandings of the nature of science. In W. McComas (Ed.), *The Nature of Science in Science Education* (pp. 83–126). Springer.

Lee, S., & Kim, H.-B. (2014). Exploring secondary students' epistemological features depending on the evaluation levels of the group model on blood circulation. *Science & Education*, *23*(5), 1075–1099.

Lehrer, R., & Schauble, L. (2006). Scientific thinking and science literacy. *Handbook of Child Psychology*, *6*, 153–196.

Leighton, J., & Gierl, M. (2007). *Cognitive diagnostic assessment for education: Theory and applications*. Cambridge University Press.

Lenhard, W., & Lenhard, A. (2016). *Computation of effect sizes*. Retrieved from https://www.psychometrica.de/effect_size.html

Liepmann, D., Beauducel, A., Brocke, B., & Amthauer, R. (2007). *Intelligenz-Struktur-Test 2000 R: IST 2000 R*. Hogrefe.

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.

Louca, L. T., & Zacharia, Z. C. (2015). Examining learning through modeling in k-6 science education. *Journal of Science Education and Technology*, *24*(2), 192–215.

Mahr, B. (2011). On the epistemology of models. In G. Abel & J. Conant (Eds.), *Berlin studies in knowledge research: Rethinking epistemology* (Vol. 1, 1st ed., pp. 301–352). De Gruyter.

Meister, S., Krell, M., Göhner, M., & Upmeier zu Belzen, A. (2021). Pre-service biology teachers' responses to first-hand anomalous data during modelling processes. *Research in Science Education*. *51*, 1459–1479. https://doi.org/10.1007/s11165-020-09929-7

Nehring, A., Nowak, K. H., Upmeier zu Belzen, A., & Tiemann, R. (2015). Predicting students' skills in the context of scientific inquiry with cognitive, motivational, and sociodemographic variables. *International Journal of Science Education*, *37*(9), 1343–1363.

Neumann, K., Härtig, H., Harms, U., & Parchmann, I. (2017). Science teacher preparation in Germany. In J. E. Pedersen, T. Isozaki, & T. Hirano (Eds.), *Model science teacher preparation programs: An international comparison of what works*. Information Age.

Newman, M. E. J. (2010). *Networks: An introduction*. Oxford University Press.

NGSS Lead States (2013). Next generation science standards: For states, by states: Appendix F - science and engineering practices in the NGSS. In NGSS Lead States (Ed.), Next generation science standards: For states, by states. The National Academies Press.

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.

Nielsen, S. S., & Nielsen, J. A. (2021a). 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

Nielsen, S. S., & Nielsen, J. A. (2021b). Models and modelling: Science teachers' perceived practice and rationales in lower secondary school in the context of a revised competence-oriented curriculum. *Eurasia Journal of Mathematics, Science and Technology Education*, *17*(4) 1–18. https://doi.org/10.29333/ejmste/10790

NRC. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press.

Oh, P. S., & Oh, S. J. (2011). What teachers of science need to know about models: An overview. *International Journal of Science Education*, *33*(8), 1109–1130.

Osborne, J. (2014). Teaching scientific practices: Meeting the challenge of change. *Journal of Science Teacher Education*, *25*(2), 177–196.

Passmore, C. M., Gouvea, J. S., & Giere, R. (2014). Models in science and in learning science: Focusing scientific practice on sense-making. In M. R. Matthews (Ed.), *International handbook of research in history, philosophy and science teaching* (pp. 1171–1202). Springer.

Passmore, C. M., & Svoboda, J. (2012). Exploring opportunities for argumentation in modeling classrooms. *International Journal of Science Education*, *34*(10), 1535–1554.

Patton, M. (1990). Purposeful sampling. *Qualitative Evaluation and Research Methods*, *2*, 169–186.

Pluta, W. J., Chinn, C. A., & Duncan, R. G. (2011). Learners' epistemic criteria for good scientific models. *Journal of Research in Science Teaching*, *48*, 486–511.

Porter, M. A., Onnela, J.-P., & Mucha, P. J. (2009). Communities in networks. *Notices of the American Mathematical Society*, *56*(9), 1082–1097.

Ronqui, J. R. F., & Travieso, G. (2015). Analyzing complex networks through correlations in centrality measurements. *Journal of Statistical Mechanics: Theory and Experiment*, *2015*(5), P05030.

Ruppert, J., Duncan, R. G., & Chinn, C. A. (2017). Disentangling the role of domain-specific knowledge in student modeling. *Research in Science Education*. *49*, 921–948. https://doi.org/10.1007/s11165-017-9656-9

Schreier, M. (2012). *Qualitative content analysis in practice*. SAGE Publications.

Schwarz, C. V. (2002). Is there a connection? The role of meta-modeling knowledge in learning with models. Retrieved from http://schwarz.wiki.educ.msu.edu/file/view/Schwarz_ICLS_metapaper.pdf

Schwarz, C. V., Reiser, B., Davis, E., Kenyon, L., Acher, 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.

Schwarz, C. V., & White, B. (2005). Metamodeling knowledge: Developing students' understanding of scientific modeling. *Cognition and Instruction*, *23*(2), 165–205.

Shi, F., Wang, L., Liu, X., & Chiu, M.-H. (2021). Development and validation of an observation protocol for measuring science teachers' modeling-based teaching performance. *Journal of Research in Science Teaching*. *58*(9), 1359–1388. https://doi.org/10.1002/tea.21712

Sikorski, T.-R. (2019). Context-dependent "Upper Anchors" for learning progressions. *Science & Education*, *28*(8), 957–981.

Sins, P. H., Savelsbergh, E. R., & van Joolingen, W. R. (2005). The difficult process of scientific modelling: An analysis of novices' reasoning during computer-based modelling. *International Journal of Science Education*, *27*(14), 1695–1721.

Sins, P. H., Savelsbergh, E. R., van Joolingen, W. R., & van Hout-Wolters, B. H. (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.

Torres, J., & Vasconcelos, C. (2015). Nature of science and models: Comparing Portuguese prospective teachers' views. *EURASIA Journal of Mathematics, Science and Technology Education*, *11*(6), 1473–1494.

Upmeier zu Belzen, A., Krüger, D., & van Driel, J. H. (Eds.) (2019). Models and modeling in science education. In *Towards a competence-based view on models and modeling in science education*. Springer International Publishing.

Upmeier zu Belzen, A., van Driel, J. H., & Krüger, D. (2019). Introducing a framework for modeling competence. In A. Upmeier zu Belzen, D. Krüger, & J. H. van Driel (Eds.), *Towards a competence-based view on models and modeling in science education* (pp. 3–19). Springer International Publishing.

VCAA (2016). Victorian curriculum: F-10. Retrieved from https://victoriancurriculum.vcaa.vic.edu.au/science/curriculum/f-10

Vo, T., Forbes, C. T., Zangori, L., & Schwarz, C. (2015). Fostering 3rd-grade students' use of scientific models with the water cycle: Elementary teachers' conceptions and practices. *International Journal of Science Education*, *37*(15), 2411–2432.

Vo, T., Forbes, C. T., Zangori, L., & Schwarz, C. (2019). A longitudinal investigation of primary inservice teachers' modelling the hydrological phenomena. *International Journal of Science Education*, *41*(18), 2788–2807.

Windschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. *Science Education*, *92*(5), 941–967.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: 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