Fostering students' meta-modelling knowledge regarding biological concept-process models

Susanne Jansen

Review committee:
Prof. dr. P.H.M. Drijvers
Prof. dr. S.E. McKenney
Prof. dr. M. Ropohl
Prof. dr. J. van Tartwijk
Dr. M.H.J. Ummels

Susanne Jansen

Fostering students' meta-modelling knowledge regarding biological concept-process models / S. Jansen / Utrecht: Freudenthal Institute, Faculty of Science, Utrecht University / FI Scientific library (formerly published as CD-ß Scientific Library), no. 115, 2022.

Dissertation Utrecht University. With references. Met een samenvatting in het Nederlands.

ISBN: 978-90-70786-54-0
Cover design: & Lay-out: Publiss | www.publiss.nl
Print: Ridderprint | www.ridderprint.nl

Fostering students' meta-modelling knowledge regarding biological concept-process models

Het ondersteunen van de *meta-modelling knowledge* van leerlingen met betrekking tot biologische concept-proces modellen

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

maandag 16 mei 2022 des ochtends te 10.15 uur

door

Susanne Jansen

geboren op 24 maart 1986
te Schiedam

Promotor:

Prof. dr. W.R. van Joolingen

Copromotor:

Dr. M.C.P.J. Knippels

Dit proefschrift werd (mede) mogelijk gemaakt met financiële steun van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) [projectnummer: 023.007.065].

Contents

CHAPTER 1

General introduction

1.1. Introduction

Models play a major role in scientific research, where scientists use them to simplify, describe and communicate complex phenomena (Svoboda & Passmore, 2013), formulate hypotheses about these phenomena, or conduct experiments (Odenbaugh, 2005). The usefulness of models is not limited to scientists only. Science and society are inherently intertwined, where results of scientific research in the form of scientific models often find their way into society. As I am writing this introduction, we are in the middle of the COVID-19 pandemic outbreak, where the news and social media all over the world show models related to contamination, herd immunity and the effect of vaccines. The ability to understand and value this kind of scientific information and models is considered to be part of scientific literacy. Several authors see scientific literacy as important for all citizens (Feinstein, 2011; Grosslight, Unger, Jay, & Smith, 1991; Oh & Oh, 2011; Schwarz et al., 2009; Sharon & Baram-Tsabari, 2020). Scientific literacy encompasses the skills that are required for understanding science in everyday life, solving problems with a scientific nature and making personal decisions on socio-scientific issues. An example of a socio-scientific issue is deciding whether or not you take a vaccine against a known disease (Lundström, Ekborg, & Ideland, 2012; Roberts & Gott, 2010).

In order for people to understand the scientific models that they encounter in daily life, knowledge about models, the creation of models and the use of models (i.e., meta-modelling knowledge) is required. This meta-modelling knowledge is part of scientific literacy (Grosslight et al., 1991). In order to stimulate citizens' meta-modelling knowledge, modelling as a scientific practice is part of the curriculum in many countries, for instance the United States Next Generation Science Standards (NGSS Lead States, 2013), the science curriculum in the Netherlands for the subjects physics, chemistry and biology (CvTE, 2018) and the National science curriculum in England (GOV.UK, 2015).

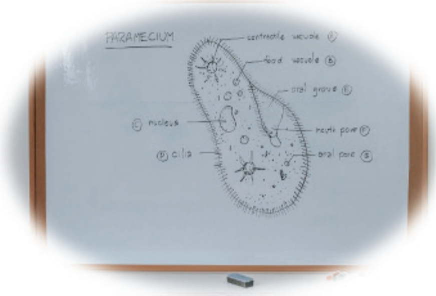
1.2. Types of models in biology education

Models can be classified into many different categories. Examples are concrete material models, verbal models, visual models, and mathematical models (Figure 1.1). In this dissertation we focus on models that teachers and textbooks use to represent biological processes and phenomena. Harrison and Treagust (2000) propose a typology of models, where the use of models ranges from concrete scale models to more abstract models such as diagrams and so called 'concept-process models'. A biological example of a scale model is a model of a torso showing the anatomy of human organs. Scale models reflect physical characteristics of objects, such as the relative size of the organs in the torso and are usually static. In abstract biological models such as diagrams and concept-process models, the concept that is referred to is not an object (such as an organ), but a biological process. A concept-process model can show the interaction of aspects that are considered important for the biological process at hand. An example of a biological concept-process model showing the light reaction of photosynthesis is depicted in Figure 1.2, where the connection between molecules, electrons, proteins, and light is visualized. The model shows how the energy from light is used by plants to generate Adenosine Triphosphate (ATP), a molecule that stores energy and is used by the plant in the formation of glucose molecules.

1.1a

1.1b

1.1c



1.1d

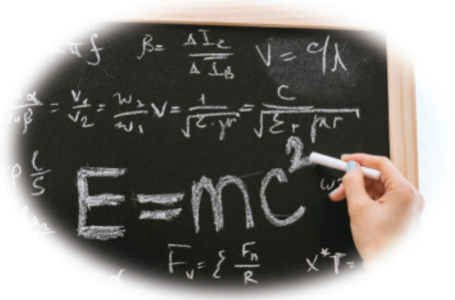


Figure 1.1. Examples of categories of models as often used in education. Figure 1.1a shows a material model of a human skeleton. Figure 1.1b shows a verbal model, where the connection of words is how relationships between aspects of a phenomenon are addressed. Figure 1.1c is an example of a visual model, showing a drawing of a 'paramecium'. Figure 1.1d is an example of a mathematical model, showing a well-known formula (all figures: Pexels.com).

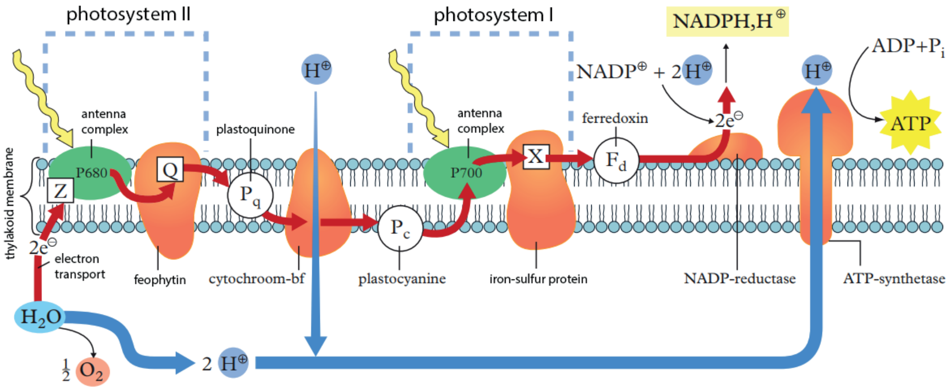


Figure 1.2. A concept-process model of the light reaction of photosynthesis, as used in secondary schools in the Netherlands. The light reaction is the first step in the creation of glucose, the ultimate goal of photosynthesis. Reprinted and translated with permission from Noordhoff Uitgevers, Groningen (Brouwens, de Groot, & Kranendonk, 2013).

Because of these visualized interactions, concept-process models cannot only be used to describe and simplify real-world situations, but also to formulate hypotheses and predict future events. Consequently, these models can be viewed upon as a suitable option for teachers to stimulate their students' meta-modelling knowledge and teach them about the nature of science.

1.3. The use of models in biology education

Even though models are an important factor when teaching about scientific practices and the nature of science, most teachers only use models as an aid in teaching science content, neglecting scientific processes such as the creation and evaluation of models, or the use of models to formulate hypotheses and predict future events (Rosária, Justi & Gilbert, 2002; Krell, Upmeier zu Belzen, & Krüger, 2012; Windschitl, Thompson, & Braaten, 2008). Passmore, Gouvea and Giere (2014) show that both students and teachers mainly view models as representations, depicting what a certain concept or phenomenon looks like or how it functions, and lack an understanding of models beyond this visualization, such as how a model can be used to formulate hypotheses. Fortus, Shwartz and Rosenfeld (2016) argue that STEM curricula and textbooks cannot be seen as a solution to this problem, since they often provide no explanation of what a model is, what the process of modelling entails or why meta-modelling knowledge is important when working with models. Most students are therefore not explicitly taught the required meta-modelling knowledge to develop an understanding of models as they are used in science.

1.4. Aims and research question

In this dissertation we aim to investigate secondary school students' meta-modelling knowledge considering biological concept-process models and develop teaching and learning activities to foster model-based reasoning with this specific type of biological models. The guiding research question for this dissertation was:

> *How can secondary school students' meta-modelling knowledge considering biological concept-process models be fostered?*

We took three steps to answer this research question:

1. Define an instrument to assess students' meta-modelling knowledge considering biological concept-process models,
2. Develop teaching and learning activities to support students' meta-modelling knowledge,
3. Explore whether students are able to apply the meta-modelling knowledge they gained to a combination of two-dimensional models and a VR model.

Concerning the **first step**, the goal was to design a specific test for biological concept-process models that would highlight multiple important aspects of model-based reasoning. This raised a few questions: which important aspects of meta-modelling knowledge can be defined and how do they apply to biological concept-process models? And when we know which aspects are important, how can we define and assess levels of reasoning for each of these aspects? Much research has previously been done on defining important aspects of meta-modelling knowledge, resulting in multiple frameworks defining various aspects and levels of reasoning (e.g., Crawford & Cullin, 2005; Grosslight et al., 1991; Justi & Gilbert, 2003; Louca, Zacharia, & Constantinou, 2011; Schwarz et al., 2009; Upmeier zu Belzen, van Driel, & Krüger, 2019). Therefore, before designing an assessment instrument, we first chose a suitable framework and tested its applicability considering biological concept-process models. Using interviews with students from various schools in the Netherlands we decided on how to adapt one of the frameworks to describe Dutch secondary school students' meta-modelling knowledge on biological concept-process models. Secondly, in order to scale up the assessments, we decided to use a test for assessing meta-modelling knowledge. We explored if and how existing tests to assess students' meta-modelling knowledge could be adapted to assess this knowledge considering biological concept-process models specifically. Eventually we based our test on the framework described by Grünkorn, Upmeier zu Belzen and Krüger (2014). After adapting and evaluating the test, we distributed the test to various schools throughout the Netherlands to assess Dutch secondary school students' meta-modelling knowledge.

After focusing on the current state of students' meta-modelling knowledge on biological concept-process models, we moved on to the **second step** of our research: designing teaching and learning activities to support students' meta-modelling knowledge considering biological concept-process models. To make sure that the developed activities both included necessary theory relating to meta-modelling knowledge and fitted the students' needs and capabilities, we needed a design approach that paid attention to both of these aspects. Multiple versions of research approaches focusing on the design of teaching and learning activities while incorporating a strong theoretical component exist (Bakker, 2018). In the spectrum of Design Research, Lesson Study (LS) has a strong focus on student learning and on the involvement of teachers in the design process of lessons. In a LS approach a team of teachers works together to design, teach, observe and evaluate lessons that are based on both literature and practice (Fernandez & Yoshida, 2004). Using LS, we first developed a lesson to introduce students to important aspects of meta-modelling knowledge, after which we developed two more lessons to support students' meta-modelling knowledge considering biological concept-process models. To make sure the theoretical component was well represented in these lessons, we differed from the usual LS approach by adding researchers to the team of teachers that designed the lessons. The goal for all three lessons was to combine the teachers' pedagogical and didactical knowledge with theory on model-based reasoning and important aspects of meta-modelling knowledge. The adapted LS approach that we used to design the lessons was evaluated after the development of the first lesson, and the test we developed during the first step of our research was used to study the effect of the three developed lessons on students' meta-modelling knowledge.

During the first two steps of our research, we let students work with two-dimensional biological concept-process models, such as the one depicted in Figure 1.1. This choice seemed logical, since the models that are present in students' textbooks and on their exams are two-dimensional representations of biological phenomena. However, three dimensional digital models have recently started to enter the classroom in the form of virtual reality (VR) environments (Velev & Zlateva, 2017). VR delivers an immersive experience, where a three-dimensional computer-generated virtual environment is created and the user is able to interact with this environment. Since the use of VR is becoming more frequent in biology education, the **third step** in our research was to explore whether students were able to apply the meta-modelling knowledge they gained while working with the activities that were developed in the second step of our research to a VR biological concept-process model. Together with Nanyang Technological University (NTU) in Singapore we developed a VR application that introduces the biological process blood-glucose regulation. The developed application showed the same information and levels of biological organization as present in students' textbooks on this topic. Students worked with the VR application, after which they compared two-dimensional models with the VR model and applied the important aspects of meta-modelling knowledge on this combination of models. The student material was used to investigate whether students were able to apply these aspects on both two-dimensional models and the VR model. The test that was developed in the first step of our research was used to find out whether working with a combination of two-dimensional models and the VR model influenced students' expressed meta-modelling knowledge.

1.5. Dissertation outline

The three steps as described in Section 1.3 are discussed in separate chapters of this dissertation. We now outline how these steps align with the studies we carried out to answer our research question. A summary of the steps and corresponding methods and participants is visualised in Figure 1.3.

In **Chapter 2** we address how we tested the applicability of the framework as described by Grünkorn et al. (2014) to biological concept-process models. This chapter describes a study in which 40 Dutch eleventh-grade students with a major in biology from four different schools in the Netherlands were interviewed to evaluate the applicability of the framework. It addresses the following research question:

> *2.1: To what extent can the described framework be used to assess students' understanding of biological concept-process models?*

The interviews were coded using the categories as described in the framework. For student answers that related to the aspects within the framework, but did not match the categories present in the framework, an addition to the framework was formulated.

In **Chapter 3** we discuss the development and application of a test to assess students' meta-modelling knowledge considering biological concept-process models. Literature showed that the context of a model can influence students' expressed meta-modelling knowledge. For example, Krell (2019) showed that the presence of a purpose of a model influences students' reasoning with the model. Al-Balushi (2011) and Krell et al. (2012) argue that the level of abstractness of a model relates to how students experience the model. The test therefore specifically focuses on 1) the effect of different types of explicit modelling-purposes on students' expressed meta-modelling knowledge, and 2) the difference in students' expressed meta-modelling knowledge considering (sub)microscopic biological concept-process models and macroscopic biological concept-process models. After distributing the test, 387 completed tests could be included in this study. The study addresses the following two research questions:

> *3.1: How does students' expressed meta-modelling knowledge related to the aspects nature of models and multiple models depend on the presence of different kinds of explicit modelling-purposes?*
>
> *3.2: How does students' expressed meta-modelling knowledge related to the aspects nature of models and multiple models differ between contexts showing macroscopic and (sub)microscopic concept-process models?*

Before developing a complete lesson series to support students' meta-modelling knowledge on biological concept-process models, a lesson was developed that served two goals: to introduce students to important aspects of meta-modelling knowledge and to investigate whether the adapted version of LS - where teachers and researchers together design a lesson - can be used as a research approach by the educational research community. Thirty-four eleventh-grade pre-university students engaged in all scheduled activities related to this study. **Chapter 4** discusses the development of this lesson and the evaluation of LS as a research approach by addressing the following two research questions:

> *4.1: To what extend does the developed lesson successfully familiarize students with important levels and aspects that are associated with model-based reasoning?*

4.2: How do teachers and researchers experience using LS as a research approach?

After evaluating the use of LS as a research approach we used the adapted version of LS to develop two lessons to support students' meta-modelling knowledge. According to Quillin and Thomas (2015), the skill to read and write visual or symbolic language, called *visual literacy*, needs to be stimulated before reasoning with biological models can be addressed. In **Chapter 5** the development of the two lessons – addressing visual literacy (Lesson 1) and model-based reasoning (Lesson 2) is discussed, including the influence of these two lessons on students' expressed meta-modelling knowledge. Two biology classes with a total of 61 eleventh grade pre-university students received these lessons. The same biology classes as the ones that were used for the study as described in Chapter 4 were used for the study described in Chapter 5. This means that students received both the lesson as described in Chapter 4 and the lessons as described in Chapter 5. The influence of the combination of these three lessons on students' expressed meta-modelling knowledge was assessed using the test as described in Chapter 3. The results of this assessment are described in Chapter 5. Figure 1.3 shows how the studies described in Chapter 4 and Chapter 5 are related. The following two research questions are addressed in Chapter 5:

5.1: What key activities are developed by the LS-team to foster students' visual literacy and engage students in modelling processes with biological concept-process models when following the suggested interventions by Quillin and Thomas (2015)?

5.2: What is the influence of the developed key activities on students' reasoning with biological concept-process models?

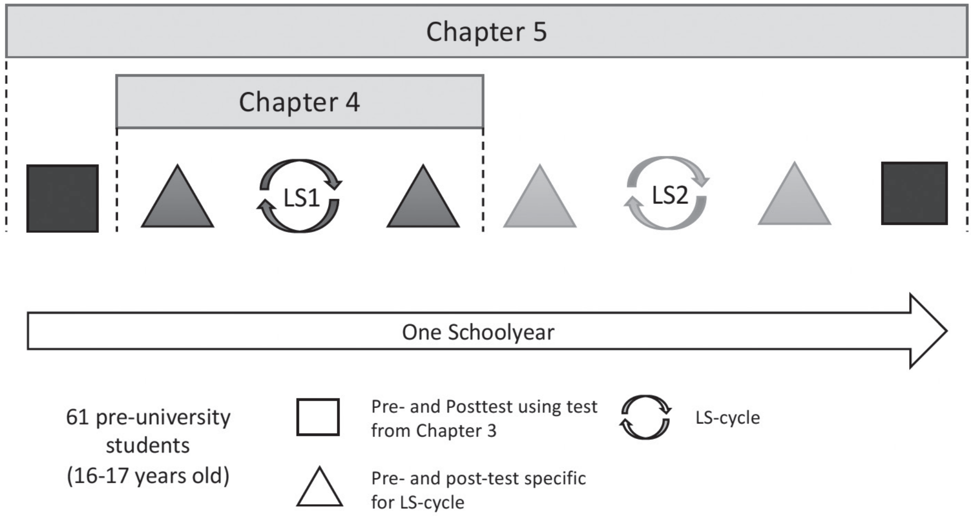


Figure 1.3. Visualisation of how the studies in Chapter 4 and Chapter 5 are related. Chapter 4 describes de first LS-cycle and corresponding pre- and post-test. Chapter 5 describes both the second LS-cycle with corresponding pre- and post-test, and the influence of both LS-cycles on students' expressed meta-modelling knowledge. The pre- and post-tests surrounding the LS-cycles (triangles) are specifically designed for each individual LS-cycle. For the pre-and post-test covering both LS-cycles (squares) we used the test as described in Chapter 3.

In our final study, discussed in **Chapter 6**, we focus on whether students are able to apply their gained meta-modelling knowledge on a combination of two-dimensional models and the VR model that we developed. To find out whether previously having worked with aspects of meta-modelling knowledge influences students' expressed meta-modelling knowledge, we applied this intervention to students who worked with these aspects during their previous academic year (studies described in Chapters 4 and 5, 41 twelfth grade pre-university students), and to a control group with students who never explicitly worked with these aspects before (47 twelfth grade pre-university students). The following research question is addressed in this study:

> *6.1: Does prior instruction in meta-modelling result in differences in applying meta-modelling knowledge to a combination of 2D models and a 3D virtual environment?*

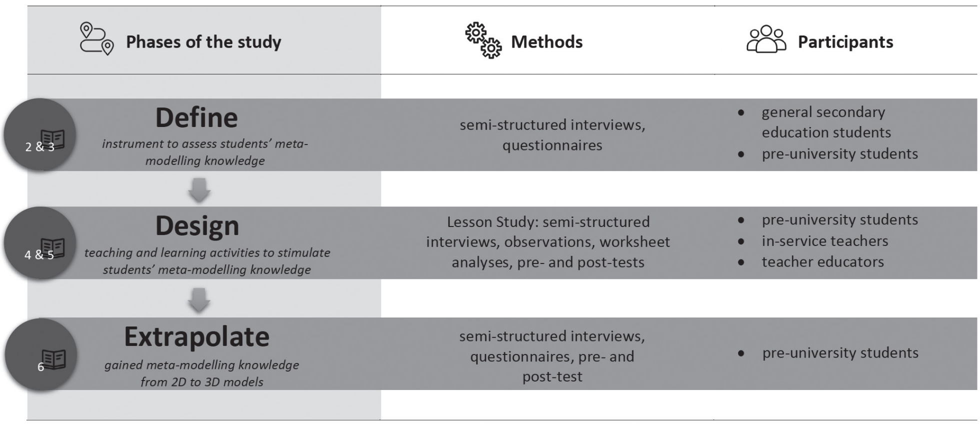


Figure 1.4. Outline of the phases of the study, showing the chapters in which each phase is discussed, the methods that were used, and the type of participants for each phase.

The main findings of all studies are summarized and interpreted in **Chapter 7**. This final chapter also discusses this dissertation's contributions, limitations, implications and directions for future research. An outline of the phases of the study as described in this dissertation is pictured in Figure 1.4.

CHAPTER 2

Assessing students' understanding of models of biological processes: a revised framework

This chapter is based on:

Jansen, S., Knippels, M.C.P.J., van Joolingen, W.R. (2019). Assessing students' understanding of models of biological processes: a revised framework. *International Journal of Science Education, 41*(8), 981-994. https://doi.org/10.1080/09500693.2019.1582821

Abstract

Models are very important tools when learning and communicating about science. Models used in secondary school biology education range from concrete scale models, such as a model of a skeleton, to abstract concept-process models, such as a visualization of meiosis. Understanding these concept-process models requires a profound understanding of the concept of models and how they are used in biology. This study evaluates the framework from Grünkorn et al. (2014) for its use in assessing students' understanding of biological concept-process models. Four additions were required to extend the applicability of the framework to concept-process models. We were also able to give an indication of students' current level of understanding of these models, showing room for improvement in all aspects of understanding. Since concept-process models have a central place in many scientific disciplines, it is important that students have a deep understanding of the nature, application and limitations of these models. The current study contributes to assessing the way students reason with concept-process models. Knowing how to improve students' view on the use of concept-process models in biology may lead to higher scientific literacy.

2.1. Introduction

Models play a key role in scientific explanations and understanding. They are representations that can describe or simplify complex phenomena, be used to make predictions about future events, facilitate the communication of ideas (Svoboda & Passmore, 2013) or make entities visible that are impossible to observe with the naked eye (Francoeur, 1997). The idea of a model is not to reach ultimate truth, but to make sense of the world. This means that they can be seen as theories, constructed to give meaning to observed phenomena (Duschl, 1990). One could say that models can either be used as representations *of something,* or that they can be used *for something*, thereby generating new data (Gouvea & Passmore, 2017). When models are used as representations, they often have a retrospective view because they summarize or visualize existing data. When models are used as a tool to predict future events and make sense of the world around us, they have a prospecting view (Krell & Krüger, 2017; Passmore et al., 2014; Upmeier zu Belzen & Krüger, 2010). By using models in both a retrospective and prospective way, they can function as a bridge between scientific theory and the world as experienced (Justi & Gilbert, 2003). Models are also present in our everyday life, such as weather models that are presented in the news, models about climate change and models about vaccination and herd immunity. Therefore, knowledge about models and the process of modelling is not only important for scientists. Having a basic understanding of models and their use in science is important for decision making processes that everyone has to go through in daily life (Odenbaugh, 2005; Oh & Oh, 2011).

Students who have an elaborate understanding of science are believed to understand that both scientific knowledge and models are human constructs, designed to explain and predict parts of phenomena (Schwarz & White, 2005). It is important for students to understand both the retrospective and prospective view on modelling. The understanding of scientific models and the process of modelling can not only facilitate the learning of science content, but also aid in developing an understanding of the way science depicts and investigates phenomena. Students' knowledge about science grows when they are able to define what has been depicted in a model and are able to relate the choices that were made when creating the model to its use. Understanding the use of models in science is part of scientific meta-knowledge (White, Collins, & Frederiksen, 2011) and adds to students' scientific literacy (Gilbert, 1991).

In secondary school biology education, the use of models ranges from concrete scale models to abstract concept-process models. The scale models, such as a model of the human skeleton, can be used to describe and simplify phenomena by depicting superficial features. The abstract concept-process models represent complex scientific processes, such as the interaction of hormones, organs and cells in the formation of sperm cells, and can be used to both make abstract entities visible and make predictions about future events. However, use of models in biology education is often limited to illustrative or communicative purposes, neglecting the scientific practice that accompanies the use of models in science. This is partly because many teachers lack experience with scientific modelling. They view models as a helpful tool for teaching about science content, but not about the nature of science (Windschitl et al., 2008). This makes it difficult for students to develop a more sophisticated scientific view on the use of models in biology.

In the learning process of students, concept-process models can be particularly useful. They are process thinking models for understanding and applying important concepts, and are considered to be the most complex type of models (Harrison & Treagust, 2000). When these models are being used in biology, they often contain positive or negative feedback loops where depending on the type of stimulus a certain response is being provoked. For example, when we look at concept-process models of the formation of sperm cells (spermatogenesis), a negative feedback loop makes sure that the production of testosterone is inhibited when its level is above a certain value. In concept-process models about this subject, several concepts such as hormones and sperm cells are connected by arrows that represent a part of the process (see Figure 2.1 below for an example of a model of this process). Because of their abstract nature, these arrows showing the effect of time and movement on a process can be difficult for students to understand. These dynamics of time and movement are specific for concept-process models. It shows that not everything that is represented is always happening at the same time or visible in real life. Whether or not the next step within a depicted process will happen depends on the factors, or concepts, that are present in a specific situation. This situation may vary depending on time and place. For example, when we look at the model of spermatogenesis, we can see that the hormones luteinizing hormone (LH) and follicle stimulating hormone (FSH) are only being released if and after the hypothalamus secreted releasing hormone (RH). When the hypothalamus is being inhibited by oestrogen and testosterone, LH and FSH will not be released. This means that one could look at these concept-process models as a collection of possible events which occur depending on the presence or absence of certain factors.

In scientific practices, such as generating and testing hypotheses about natural processes, concept-process models play a key role. Therefore, such models can be assumed to be helpful in teaching about the nature of science. However, as with other types of models, teachers almost always use concept-process models only as an illustration of a concept or a process (Windschitl et al., 2008), missing the opportunity for students to learn about scientific practices. To get students to develop a more sophisticated scientific view on the use of models and be able to improve the way they reason with biological concept-process models, a better insight in students' understanding of these models is necessary.

To assess students' understanding of models that are being used in science education, various frameworks have been developed. Where some frameworks focus on practical skills when it comes to modelling (e.g., Louca et al., 2011; Schwarz et al., 2009), others focus on learners' understanding of models and their use in science (e.g., Crawford & Cullin, 2005; Grosslight et al., 1991; Justi & Gilbert, 2003). Upmeier zu Belzen and Krüger (2010) combined the frameworks from Crawford and Cullin (2005), Grosslight et al. (1991) and Justi and Gilbert (2003) with theoretical approaches from Mahr (2009) and Giere (2004). This resulted in a theoretical framework for assessing and investigating learners' understanding of models and their use in science. The framework contains 5 aspects concerning model understanding: *nature of models, multiple models*, *purpose of models, testing models* and *changing models*. In line with other researchers (Crawford & Cullin, 2005; Grosslight et al., 1991; Schwarz et al., 2009) several levels of understanding have been formulated. However, when formulating the different levels of understanding, Upmeier zu Belzen and Krüger (2010) explicitly refer to Mahr, (2009), who defines models based on both the creation and the application of a model. This way the framework addresses both the understanding of models as an illustration of something and the understanding of models from a more methodological perspective: as an instrument to test ideas or draw conclusions while reflecting the scientific perspective of models as research tools (Treagust, Chittleborough, & Mamiala, 2004). To validate the theoretical framework, Grünkorn et al. (2014) have tested its use in assessing students' understanding of biological models.

The five aspects within the framework reflect different characteristics of the understanding of models and their use in science. The aspects *nature of models* and *multiple models* focus on ontological and epistemological concepts of models. These two aspects reflect on the way models can describe and simplify phenomena *(model of something)*. The aspects *purpose of models*, *testing models* and *changing models* focus on cognitive processes and the way these models are used in science. These three aspects together reflect on the fact that models can be used to test hypotheses, make predictions about future events and communicate ideas *(model for something)* (Grosslight et al., 1991; Justi & Gilbert, 2003). The aspects can also be related to the cognitive model of science as described by Izquierdo-Aymerich and Adúriz-Bravo (2003), which focuses on how scientists do science using models. Scientists choose a strategy to pursue a scientific goal and select a model that is most appropriate for reaching this goal. The aspects *purpose of models* can be related to the goal that a researcher is trying to achieve, while the aspects *testing models* and *changing models* reflect on the selection of an appropriate model to achieve this goal.

In the aspect *nature of models*, students compare the model with the original and explain the extent to which the model can be compared with the original. The aspect *multiple models* refers to the fact that multiple models can be used to represent the same original, focussing on different aspects of the original or reflecting on them in a different way. Since no single model can illustrate an object or process in all its aspects, models are used to explore important and difficult aspects of a concept. For this, the model is often simplified, fitting the questioner's need and prior knowledge and emphasizing only those aspects that are of great importance to explain a certain key idea (Harrison & Treagust, 2000). With the aspect *purpose of models,* the framework tries to differentiate in the way students think about the purpose of models in science and whether the models can be used to predict the outcome of a certain input. The aspect *testing models* describes the way students think a model is validated, and the aspect *changing models* reflects on the fact that models are by definition changeable and temporary. For example, when falsification of a hypothesis about the original calls for changes in the currently used model. For all of these aspects, up to four levels of students' understandings were determined, ranging from a low level of understanding to an expert level of understanding (Grünkorn et al., 2014).

Since concept-process models can be particularly useful when learning about scientific reasoning, but are also conceived as the most difficult models for students to reason with, it is important to know secondary students' level of understanding for this type of models. Because the framework from Grünkorn et al. (2014) has been tested for its use in assessing students' understanding of biological models, it seems the right fit for assessing students' understanding of biological concept-process models. However, to our knowledge, this framework has not specifically been tested for its use in assessing students' understanding of this type of biological models.

We aim to evaluate the usability of the framework from Grünkorn et al., (2014) - from this point referred to as 'Grünkorn's framework'- for the assessment of students' understanding of biological concept-process models. Therefore, the research question for this study is:

> *2.1: To what extent can the described framework be used to assess students' understanding of biological concept-process models?*

The aspects and levels within the framework have been validated in a previous study (Grünkorn et al., 2014). This study focuses on determining whether the given categories of student answers as described in Grünkorn's framework provide a valid assessment of students' understanding of biological concept-process models and whether or not these categories need adaptation in this new context of use. Also, since the data that were collected during this study consist of student answers relating to the various levels and aspects within the framework, these answers are used to give an indication of students' current understanding of biological concept-process models.

2.2. Method

To answer our research question, we interviewed students about two different concept-process models. The choice for task-based interviews was made so that we were able to get more in-depth insight in the reasons behind students' answers. This made it easier to judge whether the answers matched the existing descriptions of categories within Grünkorn's framework, or if an addition to the framework was necessary. The interview questions related to the five aspects within this framework. Student answers were then compared with the description of the levels within Grünkorn's framework to determine whether the answers matched the existing levels. Since all answers relating to the five different aspects should fit one of the categories within the framework, answers that we could not match with the existing categories resulted in an addition to this framework.

2.2.1. Participants

Forty Dutch eleventh-grade students with a major in biology (16-18 years old, 19 females and 21 males) from four schools in the Netherlands participated in this study. Schools were chosen from different areas in the country. Per school ten students participated, five on pre-university level and five on higher general education level. Students were randomly selected, except from one school where students were assigned in class by the researcher.

2.2.2. Interviews

The task-based interview scheme was based on the five aspects within Grünkorn's framework. Questions mostly related directly to one of the aspects. For example, the question 'How can be determined whether this model is correct?' relates to the aspect *testing models*. Other questions, such as 'What is the meaning of the arrows in this model?' were meant to get a deeper insight in the way students interpret different aspects or concepts that were mentioned in the model. Two pilots with a total of 6 secondary students were conducted to optimize the interview scheme. The students in the pilot studies were 16-18 years old and followed a major in biology (3 females and 3 males). A list containing all translated interview questions can be found in Appendix A.

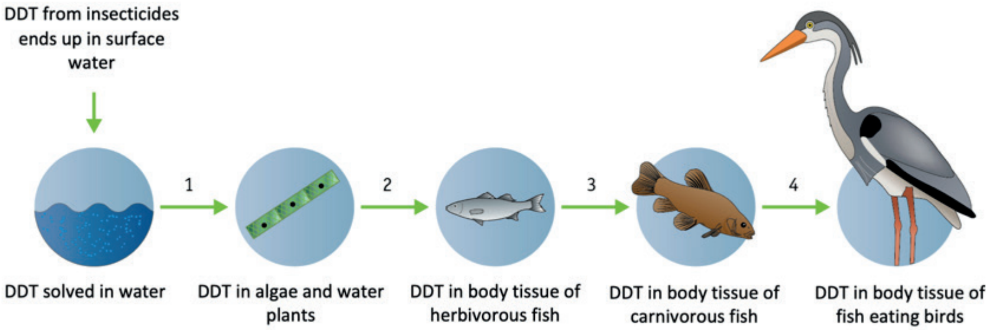


Figure 2.1. A concept-process model about the accumulation of DDT in the food chain. The model shows how a substance can accumulate in higher order organisms. Provided by and translated with permission from Biologie voor jou, Malmberg.

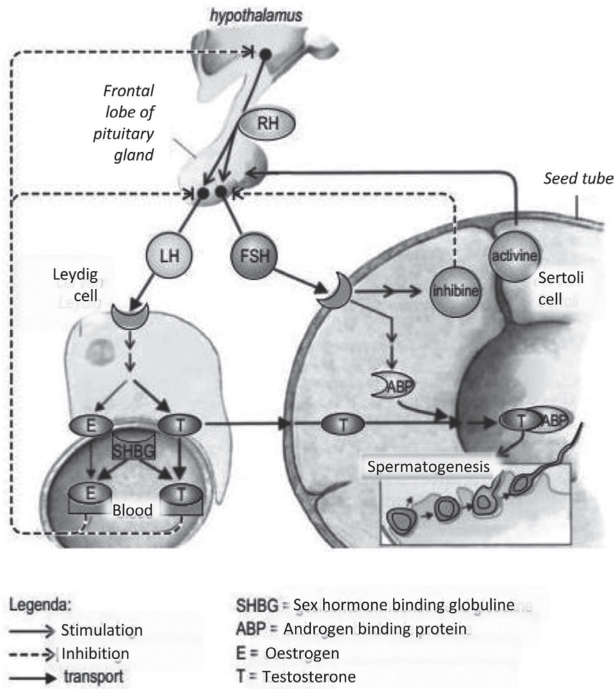


Figure 2.2. A concept-process model showing the role of hormones in the formation of sperm cells. Reprinted and translated with permission from archive CvTE (2016) [VWO Examen biologie tweede tijdvak, retrieved from https://www.examenblad.nl/examen/biologie-vwo-2/2016/vwo?topparent=vg41h1h4i9qe]

We used two concept-process models: one model showing the process of accumulation of DDT in the food chain (Figure 2.1) and one model about spermatogenesis, the formation of sperm cells (Figure 2.2). Three criteria were used for selecting these models. First, we selected models that used arrows to indicate the dynamics of a certain process. This way both models are concept-process models that use the same type of indicator for the effect of time and movement. Second, we chose models that are not present in schoolbooks to avoid a difference in pre-knowledge concerning how to reason with a specific model. Third, we chose two models that show processes on a different level of biological organization. As Knippels and Waarlo (2018) stress, students find it difficult to reason between different levels of biological organization. In this case the model about the accumulation of DDT is on the level of ecological community and the model about spermatogenesis shows a process on both the molecular and cellular level. By presenting both models, students had to reason with both abstract and concrete concepts.

Students were interviewed individually. The interviews took about 20-30 minutes per interview and were audiotaped. To avoid bias, interviews about the two different models were counterbalanced. Half of the students started with the model about the accumulation of DDT, the other half of the students started with the model about spermatogenesis.

2.2.3. Analysis

The interviews were transcribed verbatim and coded using categories based on Grünkorn's framework. For example, the code 'Nature II' was used to indicate that a text fragment said something about the aspect *nature of models,* and matched with the *second level* within the framework. For student answers that related to the aspects within the framework, but did not match the description of an existing level, the level of the answer was assessed and an addition to the framework was formulated. This resulted in an extension of the framework, making it suitable for application on concept-process models. The framework that was used to code the interviews is depicted in Appendix B.

Ten percent of the interviews were coded by a second independent coder. For this, the second coder received the transcribed interview in which the fragments that were coded by the first coder were highlighted. This resulted in a Cohens' Kappa of 0.73.

2.3. Results

2.3.1. Assessment of the framework

Not all student answers fit in the existing descriptions of the levels within Grünkorn's framework. This resulted in four additions to the existing framework (Table 2.1). First, since the models that were used showed processes, students often described these processes mentioning multiple concepts that were depicted in the model. In Grünkorn's framework, the second and third level of the aspect nature of models fit well with the descriptions students gave about concept-process models. However, the first level of *nature of models* as described in the framework mentions models to be exact replicas of the original. Since concept-process models often contain drawings of arrows to connect several concepts, no student mentioned that this model including the arrows exactly matched the original. Instead, they talked about the topic of the model, explaining that these concepts together represent the process that is being depicted:

> *(Student_VM3M): I think this [model] has something to do with food. It shows that the algae use the DDT as a food source. And small fish eat algae. And carnivorous fish eat the smaller fish. And then the heron, a fish-eating bird, eats the other fish.*

This way of explaining the nature of models is specific for concept-process models, since the dynamics of time and movement are incorporated in the model. Students therefore describe what the arrows mean and are not just talking about the objects or concepts like they would do for non-dynamic biological models. To highlight this difference, we added the description of the process as shown in the model to the aspect *nature of models*. Since students mentioned that the way the process functions is the same as it functions in the real world, this description of the process can be seen as a matching reference to the original and has therefore been added to the first level of the aspect *nature of models.*

Second, when looking at the aspect *purpose of models,* mentioning the different concepts that are present in the model and linking them would seem to lead to a level 2 understanding, since level 2 is defined by the ability to define relationships. However, since these links or relationships between concepts are already depicted in the model and were not thought of by these students themselves, one could say that the students were just reading out loud what the model was showing them instead of identifying relationships:

> *(Researcher): How could this model be used?*
>
> *(Student_HJ1S): I think, well, perhaps for students. To show them like, here you have producers and here you have consumers. So, you can explain that these eat these [pointing at organisms in model], and so on.*

Also, by just following the arrows the students neglected the dynamics of space and time that are specific for concept-process models, which explain the relationships between the different concepts within the model. Therefore, just describing the events that are depicted in a model was added to the first level of the aspect *purpose of models* within the framework.

Table 2.1. The revised framework, based on the framework from Grünkorn et al. (2014). The framework includes additions (depicted in bold) that make it possible to assess students' understanding of biological concept-process models.

	Initial level	Level 1	Level 2	Level 3
Nature of models		- Model as copy - Model with great similarity - Model represents a (non-) subjective conception of the original **- Displays a process, its components and how they are related**	- Parts of the model are a copy - Model as a possible variant - Model as focused representation	- Model as hypothetical representation
Purpose of models		- Model for showing the facts **- Model for showing events**	- Model to identify relationships	- Model to examine abstract/ concrete ideas
Multiple models	- All models are the same - Various models of different originals - Only one final and correct model	- Different model object properties	- Focus on different aspects	- Different assumptions
Testing models	- No testing of models **- Perceiving schoolbooks or their authors as authorities providing absolute truth**	- Testing of material - Testing of basic requirements	- Comparison between original and model - Comparison and matching of original and model	- Testing hypotheses - Testing of hypotheses with research designs
Changing models	- No reason for alterations - Alteration of how different originals are represented	- Alterations to improve the model object - Alterations when there are errors in the model object - Alterations when basic requirements are not met	- Alterations when model does not match the original - Alterations due to new findings about the original	- Alterations due to findings from model experiments **- Alterations when the focus of the model shifts to a different aspect of the process**

The third addition was made to the initial level of the aspect *testing models*. Several students mentioned that the only way to test whether or not the model was correct was to compare the model with their schoolbook or ask the authors of their schoolbooks whether the model was correct:

> *(Student_HJ1M): I think that the creator of the schoolbook, that he knows how it works.*

They gave their schoolbooks or the authors authority by viewing their information as the absolute truth. This means that the students were not really testing the model, but believed in the authority of teachers and schoolbook authors, and only checked whether the model agreed with what their schoolbook said. Therefore, we added this way of reasoning to the initial level of understanding for this aspect.

The fourth and final addition was made to the aspect *changing models*. A model can focus on a certain aspect of a process. For example, a model about spermatogenesis can focus on the role of hormones in the formation of sperm cells, but also on the distribution of chromosomes in the process of meiosis. Students mentioned changing a model to shift the focus that is being depicted in the model:

> *(Researcher): What would be a reason to change this image? And perhaps you can also add what kind of change that would be?*
>
> *(Student_HJ2E): Yes, well, when the purpose of the image is different [...] This is quite abstract I think. So, in theory, well this is mostly about the nucleus. So, you could also make clear how the rest of the cell is being influenced by this process. Or perhaps in which organs this happens [...]*

This way of reasoning requires knowledge about the fact that a model does not match the original, but is a focused representation. It also shows that the student knows that multiple stories can be told about the same process by shifting the focus of the model. Since level 2 for the aspect *changing models* only includes alterations when the model does not match the original, we decided that mentioning alterations to shift the focus to a different aspect exceeded this level of reasoning. Therefore, we added this type of alterations to the third level for this aspect. With these four additions, all utterances that referred to model understanding could be classified in one of the levels of the revised framework. This revised version of the framework is shown in Table 2.1.

2.3.2. Indication of level of understanding

In general, almost all students reached level 1 or 2 for all aspects mentioned within the framework. Almost none of the students reached the highest levels within the revised framework, except within the aspect *testing models* (Table 2.2).

Table 2.2. Percentage of students that reached a certain level within the framework (n=40). The aspects and levels are shown on the left. 'Accumulation of DDT' shows the percentages for the model on the level of ecological community and 'Spermatogenesis' shows the percentages for the model on the molecular/cellular level.

	Accumulation of DDT	**Spermatogenesis**
Changing initial	10%	15%
Changing 1	44%	33%
Changing 2	46%	45%
Changing 3	0%	6%
Multiple initial	3%	0%
Multiple 1	8%	3%
Multiple 2	83%	90%
Multiple 3	8%	8%
Nature 1	30%	3%
Nature 2	70%	95%
Nature 3	0%	3%
Purpose 1	56%	32%
Purpose 2	44%	61%
Purpose 3	0%	8%
Testing initial	26%	31%
Testing 1	5%	8%
Testing 2	67%	36%
Testing 3	3%	25%

Two types of distinctions in students' understanding can be made. First, a difference in level of understanding can be observed when looking at the two different models that were used. As shown in Table 2.2, level 3 has been reached much more often by students when asked about the model on the cellular and molecular level than when asked about the model on the level of ecological community. This especially holds for the aspect *testing models*. For the model on the cellular and molecular level, which shows the formation of sperm cells, students often talk about 'adding substances' to find out what that does to the formation of sperm cells. This way of formulating a hypothesis, based on the model that is presented, corresponds to the third level of the aspect *testing models:*

(Researcher): How has been determined whether this image is correct?

[...]

(Student_VM3M): You cannot just say something [...] It has to be clear and correct. So, I think people will actually look at it like, is this correct? [...] If there are people who want to function as subjects, then they can do tests on them. Like, if I add this substance to your body, how will your hormones respond to that?

(Researcher): So, you add something?

(Student_VM3M): Yes, an inhibiting substance for example. So that they can say, if we add this substance, this particular process will be inhibited. Does it really have that effect?

However, even though some students talked about things that might happen when adding substances or changing aspects in the given model, none of them used the word hypothesis or specifically mentioned a link between hypotheses and the use of models to test these. Instead, when talking about changing aspects within the model, they often referred to an assignment they had in class where students had to explain what would happen if a given aspect changed.

(Student_VJ1E): Oh, we had something like this on an exam! [...] They asked what the influence was of extra testosterone on the fertility of a man.

The second distinction in level of understanding can be made between students from pre-university level and students from higher general education level. Pre-university level students show a higher level of understanding for all five aspects within the revised framework (Table 2.3).

2.4. Conclusion and discussion

Concept-process models can be particularly useful for students when learning about the nature of science. This type of models is the most complex type of models, showing a process in which for example the effect of time, movement and feedback loops is visualized. These models can be helpful for students when learning about scientific practices such as testing hypotheses. However, since concept-process models in biology education are mostly used for illustrative or communicative purposes, the opportunity to let students get more acquainted with the scientific practice that accompanies the use of models in science is often missed. This makes it difficult for students to develop a more sophisticated scientific view on the use of models in biology.

Table 2.3. Percentage of students that reached a certain level within the framework (n=40). Both models are included. The aspects and levels are shown on the left. A distinction has been made between students at higher general education level and students at pre-university level.

	Higher general education level	**Pre-university level**
Changing initial	11%	13%
Changing 1	57%	23%
Changing 2	29%	62%
Changing 3	3%	3%
Multiple initial	3%	0%
Multiple 1	10%	0%
Multiple 2	79%	93%
Multiple 3	8%	7%
Nature 1	24%	10%
Nature 2	76%	88%
Nature 3	0%	2%
Purpose 1	56%	33%
Purpose 2	38%	65%
Purpose 3	5%	3%
Testing initial	42%	15%
Testing 1	6%	8%
Testing 2	42%	62%
Testing 3	11%	15%

The current study investigated to what extent Grünkorn's framework can be applied to assess students' understanding of biological concept-process models. Our results show that additions to the framework were needed. Four descriptions were added to different levels and aspects within the framework:

Nature of models, level 1: displays a process, its components and how they are related

Purpose of models, level 1: model for showing events

Testing models, initial level: perceiving schoolbooks or their authors as authorities providing absolute truth

Changing models, level 3: alterations when the focus of the model shifts to a different aspect of the process

We could not match all of the descriptions present in Grünkorn’s framework to our results. A reason for this could be that some of the descriptions simply do not fit with biological *concept-process models*. For example, ‘testing the material’ might be an appropriate answer when asked how to test if a scale model of a bird is correct. But when we look at a drawing of a process, testing the material is not a likely thing to do when you want to see if the model is correct. However, our aim was to find out if Grünkorn’s framework could be used to assess students’ understanding of concept-process models, without ruling out the assessment of other types of biological models. Therefore, the only adjustments we made to the framework were additions that were necessary for students’ understanding of biological concept-process models, leaving all the other descriptions in place.

Since we matched students’ answers to the levels within the revised framework, it is possible to give an indication of the current level of students’ understanding of biological concept-process models. We found that in general students reached level 2 for most aspects within this framework. However, two types of distinctions in students’ understanding can be made. First of all, students reached the highest levels within this framework more often when asked about the model on the molecular/cellular level than when asked about the model on the level of ecological community. This especially holds for the aspect *testing models*. Some students mentioned that a model such as the one we presented on the molecular and cellular level was used in an exam at school. Students had to use the model to explain what happened when a certain substance was added to the original: a human being. This means that they had to formulate a hypothesis, an activity that corresponds with the highest levels of reasoning within the framework. No student mentioned having done an exercise like that when they were confronted with the model on the level of ecological community. The fact that students had done such an exercise with models on the molecular and cellular level of organization could have triggered answers that reach the highest levels within the framework. However, further research is necessary to confirm whether this is the case.

Table 2.3. Percentage of students that reached a certain level within the framework (n=40). Both models are included. The aspects and levels are shown on the left. A distinction has been made between students at higher general education level and students at pre-university level.

	Higher general education level	**Pre-university level**
Changing initial	11%	13%
Changing 1	57%	23%
Changing 2	29%	62%
Changing 3	3%	3%
Multiple initial	3%	0%
Multiple 1	10%	0%
Multiple 2	79%	93%
Multiple 3	8%	7%
Nature 1	24%	10%
Nature 2	76%	88%
Nature 3	0%	2%
Purpose 1	56%	33%
Purpose 2	38%	65%
Purpose 3	5%	3%
Testing initial	42%	15%
Testing 1	6%	8%
Testing 2	42%	62%
Testing 3	11%	15%

The second distinction in levels of reasoning could be made between students from higher general education level and students from pre-university level. In general, students from pre-university level reached the highest levels within the framework more often than students from higher general education level. Students at pre-university level are being prepared for studying at university, where they are expected to be able to reason with models in a scientific way. Therefore, it is possible that students at pre-university level get confronted more often with situations that trigger reasoning on the highest levels within the framework than students at higher general education level. More research has to focus on reasons to explain this difference.

In this study we used the framework from Grünkorn et al. (2014) for assessing students' understanding of biological concept-process models and extended its applicability to these models. Since these types of models are very frequently used in biology education, an adequate understanding of and the ability to reason with concept-process models is crucial for an in-depth understanding of biological phenomena. Outside of biology such models often appear in other sciences, such as physics, chemistry and economy. Given their central place in many scientific disciplines, a deep understanding of nature, application and limitations of concept-process models is an essential part of students' scientific literacy. The current study contributes to assessing the way students reason with this type of models which may lead to higher scientific literacy and their understanding of the nature of science. A next step would be to investigate how to foster students' reasoning with concept-process models in secondary biology education in order to leverage their understanding of models of biological processes.

CHAPTER 3

Secondary school students' meta-modelling knowledge of biological concept-process models

Jansen, S., Knippels, M.C.P.J.,
van Montfort, M., & van Joolingen, W.R.

Abstract

The use of scientific models in class is a common practice. One of the more common types of models used in biology education are models to describe biological processes (e.g., concept-process models). Even though the dynamic nature of these models can be of great value to teach students about the nature of science, not much is known about students' knowledge on models or the use of models (i.e., meta-modelling knowledge) for this specific type of biological model. This study therefore aims to describe pre-university students' expressed meta-modelling knowledge considering biological concept-process models. We developed a test where we take both the effect of different types of explicit modelling-purposes and the effect of a (sub)microscopic or macroscopic context on students' expressed meta-modelling knowledge into account. In total 387 eleventh-grade pre-university students (16-18 years old) from 16 schools in the Netherlands completed the test. Two repeated measures ANOVA's were carried out to describe the effect of explicit modelling purposes and the type of model ((sub)microscopic or macroscopic) on students' expressed meta-modelling knowledge. Considering the effect of different types of modelling purposes, our results did not indicate an effect of these types of purposes on students' expressed meta-modelling knowledge considering biological concept-process models. Also, our results showed that students' expressed meta-modelling knowledge is more advanced for (sub)microscopic models than for macroscopic models. We discuss possible explanations for these differences and suggest focal points for designing teaching material that intends to foster students' meta-modelling knowledge considering biological concept-process models.

3.1. Introduction

Knowledge about the use of models in science is not only useful for scientists, but also for non-scientists. In science, models help to simplify, describe and communicate complex phenomena (Svoboda & Passmore, 2013), and provide conceptual frameworks that allow scientists to formulate hypotheses and conduct experiments (Odenbaugh, 2005). Science and society are inherently intertwined, and results of scientific research often find their way into society (e.g., models about climate change, vaccines, or the course of a pandemic outbreak). To understand and value this scientific information, citizens should be scientifically literate (Feinstein, 2011; Grosslight et al., 1991; Oh & Oh, 2011; Schwarz et al., 2009; Sharon & Baram-Tsabari, 2020). Scientific literacy encompasses the skills that are required for understanding science in everyday life and making personal decisions on socio-scientific issues, such as the vaccination debate in society (Lundström et al., 2012; Roberts & Gott, 2010). In order for people to understand the scientific models that they encounter in daily life, knowledge about models, the creation of models and the use of models (i.e., meta-modelling knowledge) is required. This meta-modelling knowledge is part of scientific literacy (Grosslight et al., 1991). For this reason, modelling as a scientific practice is a part of the curriculum in many countries (e.g., the United States Next Generation Science Standards (NGSS Lead States, 2013), the science curriculum in the Netherlands for the subjects physics, chemistry and biology (CvTE, 2018) and the National science curriculum in England GOV.UK, 2015)).

Hodson (2003) identifies four main purposes of secondary school science education: 1) learning science and technology, 2) learning about science and technology 3) doing science and technology and 4) engaging in socio-political action. Models play a role in each of these four purposes. First, models can function as a teaching aid, where they can illustrate and communicate theories. Second, models can aid in teaching about the *nature of science* and give students insight in the complexity of science and scientific practices. Third, modelling practices where students create, evaluate and revise a model, let students participate in doing science. The third and fourth purpose can be combined by engaging students in scientific problem-solving situations that have a community dimension (Hodson, 2003). An example of such a situation is the use of models in explaining the pandemic outbreak of the COVID-19 virus in 2020 and the way these models are used to influence societal behaviour. Students who understand the creation and use of models develop an understanding of how science investigates phenomena.

Even though models are an important factor when teaching about scientific practices, such as formulating hypotheses or simplifying complex phenomena, most teachers only use models as an aid in teaching science content, neglecting the scientific process that creates, predicts and evaluates these models (Justi & Gilbert, 2002; Krell et al., 2012; Windschitl et al., 2008). Passmore et al., (2014) and Upmeier zu Belzen and Krüger (2010) point out the importance of understanding both the retrospective view on modelling and the prospective view on modelling when it comes to meta-modelling knowledge. In the retrospective view, the model has a representational function, depicting what a certain concept or phenomenon looks like or how it functions (model of something). In the prospective view, the attention shifts to how the components and relations that are depicted in a model serve an epistemic purpose beyond depiction, such as testing hypotheses (model for something). Passmore et al. (2014) argue that students and teachers mainly show a retrospective view on modelling and lack a prospective view on modelling. Fortus et al. (2016) add that the STEM curriculum and textbooks they studied provide no explanation of what a model is, what the process of modelling entails or why meta-modelling knowledge is important when working with models. This means that students are therefore not explicitly taught the required meta-modelling knowledge to develop an understanding of models as they are used in science. A much discussed solution to this problem can be to incorporate suitable model-based learning approaches into the science curriculum (e.g., Gilbert & Justi, 2016; Krell et al., 2012; Quillin & Thomas, 2015; Schwarz & White, 2005; Treagust, Chittleborough, & Mamiala, 2002).

Focusing on biology education, various models are used in class to describe biological phenomena. In this study we focus on a specific type of model that is often used to describe biological processes: concept-process models (Harrison & Treagust, 2000). Even though models of biological processes can be used in class for various scientific purposes, such as formulating hypotheses and predicting future events, not much is known about students' meta-modelling knowledge considering this type of models. Therefore, this study aims to describe pre-university students' meta-modelling knowledge for biological concept-process models specifically. Knowing about students' meta-modelling knowledge can provide a basis for developing model-based learning approaches that focus on *concept-process models* specifically.

3.2. Theoretical Framework

3.2.1. Classification of models in biology education

Models can be classified according to many different criteria. For example, a semantic classification is suggested by Frigg and Hartmann (2006), who distinguish between representational models and models of theory, referring to respectively models visualizing a part of the world and theoretical structures which take all propositions of a theory into account. Boulter and Buckley (2000) put forward an ontological classification, suggesting that models can have different modes of representation. Examples are concrete material models, verbal models, visual models, mathematical models, and gestural models. According to an epistemological classification by Gilbert, Boulter and Elmer (2000), models can be categorized into mental models, expressed models, scientific models, historical models, and teaching models.

Since students' often work with textbooks in class, this study focusses on analogical models that teachers and textbooks use to represent scientific biological processes and phenomena. Harrison and Treagust (2000) propose a typology of models in which they attempt to characterize similarities and differences in the analogical teaching and learning models that are used in science education. In this typology, the use of models ranges from concrete scale models to more abstract models such as diagrams and concept-process models. A biological example of a scale model is a model of a torso showing the anatomy of human organs. Scale models reflect physical characteristics of objects, such as the relative size of the organs in the torso and are usually static. Concept-process models are abstract models, where the concept that is referred to is not an object (such as an organ), but a biological process such as photosynthesis. In a concept-process model of the process of photosynthesis, a model can show the interaction of molecules, electrons, proteins and light to create ATP, a molecule that is used to store energy and generate glucose (Figure 3.1). Because of these dynamic interactions, concept-process models can not only be used to describe and simplify the real-world situation, but also to formulate hypotheses and predict future events. Describing and simplifying the real-world situation relates to the way models are created, and formulating hypotheses is one of the key aspects of 'doing science'. Consequently, these models can be viewed upon as a suitable option for teachers to stimulate their students' meta-modelling knowledge and teach them about the nature of science. This study will focus on students' current expressed meta-modelling knowledge considering biological concept-process models.

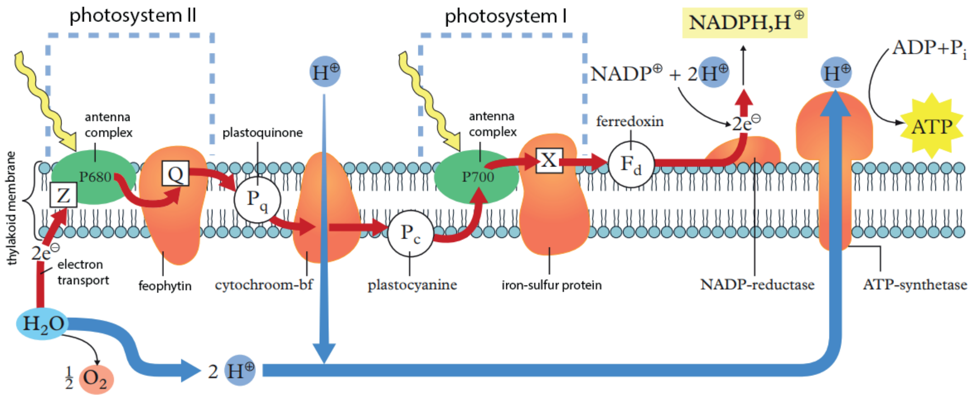


Figure 3.1. A concept-process model of the light reaction of photosynthesis, as used in secondary schools in the Netherlands. The light reaction is the first step in the creation of glucose, the ultimate goal of photosynthesis. Reprinted and translated with permission from Noordhoff Uitgevers, Groningen (Brouwens et al., 2013).

3.2.2. Framework to assess meta-modelling knowledge

Various frameworks have been developed to assess students' expressed meta-modelling knowledge, addressing aspects that are of importance when reasoning with models in science. They focus on aspects that are important to understand models and their use in science, such as the nature and purpose of models (e.g., Crawford & Cullin, 2005; Grosslight et al., 1991; Justi & Gilbert, 2003) or the handling of models, which involves more practical skills (e.g., Louca et al., 2011; Schwarz et al., 2009). Upmeier zu Belzen & Krüger (2010) combined the frameworks from Crawford & Cullin (2005), Grosslight et al. (1991) and Justi & Gilbert (2003) to create a theoretical framework that can be used as an analytical framework for assessing and investigating students' meta-modelling knowledge (Upmeier zu Belzen et al.,2019). Unlike the approach described by Schwarz et al. (2009), who included practical skills in their described learning progression, the theoretical framework by Upmeier zu Belzen & Krüger, (2010) solely emphasizes the cognitive component of models. Grünkorn et al. (2014) evaluated this theoretical framework by assessing the meta-modelling knowledge of secondary school students using different biological model contexts. Jansen, Knippels, & van Joolingen (2019) then extended the framework to include categories that are specific for biological concept-process models (Table 3.1). Since this extended framework from Jansen et al. (2019) addresses biological concept-process models specifically, we decided to use this framework as a basis to develop a test in the current study.

The assessment framework for meta-modelling knowledge consists of five aspects: (1) *nature of models*, (2) *multiple models*, (3) *purpose of models*, (4) *testing models*, and (5) *changing models*. The aspects *nature of models* and *multiple models* focus on the ontological and epistemological understanding of already existing models, while the other three aspects (*testing models, changing models* and *purpose of models*) reflect on the way models are used in science: to communicate ideas, make predictions about future events and test hypotheses. For each of these aspects, three levels of understanding are defined (Grünkorn et al., 2014). On Level 1, students think of a model as an object, without focusing on the original (*model object* (Mahr, 2009)). Students have reached Level 2 when they perceive the model as a medium of something, understand that the model is created by people and that it can be used to communicate about the original (*model of something* (Mahr, 2009)). The most advanced level of model understanding, Level 3, represents a methodological view of models in which students comprehend the use of models in science. An example of such a use is to formulate or test hypotheses about the original (*model for something* (Mahr, 2009)). Krell et al. (2012) found that these levels show an increasing degree of difficulty. Viewing these levels in the context of the research of Passmore et al. (2014), Level 1 and Level 2 can be linked to a retrospective view on modelling, while Level 3 relates to a prospective view on modelling. Next to the three levels of complexity, an initial level of understanding has been defined for the aspects *testing models, changing models* and *multiple models* (Grünkorn et al., 2014). Student responses fall in this category when they lack a basic understanding of these aspects of meta-modelling knowledge by rejecting the fact that models can be tested or changed, or that multiple models about the same topic exist. Unlike Level 1-3, the initial level of understanding cannot be seen as a valid way of reasoning.

The current study focuses on students' expressed meta-modelling knowledge considering existing biological concept-process models, without incorporating modelling practices where students would have to test or change a model. Therefore, the aspects *nature of models* and *multiple models,* which reflect on the epistemological and ontological understanding of models, will be the focus of this study. For the aspect *nature of models,* the three levels of understanding reflect on the extent to which a model can be compared to the original (Table 3.1). Considering concept-process models, students can think that the model displays a process, different components and how they are related, without linking the model to the original (Level 1); be aware of the fact that models are often simplifications, fitted to the questioners' need and prior knowledge, and therefore only focus on a part of the process (Level 2); or know that models can be used to highlight a certain hypothesis relating to the depicted process (Level 3) (see Table 3.1) (Jansen et al., 2019). The aspect *multiple models* addresses the fact that multiple models can be used to represent the same original. No single concept-process model can show every detail of a process or explain all existing hypotheses and ideas. This leads to a variety of models addressing the same process, but varying in aesthetic choices (Level 1), showing a difference in focus (Level 2), or explaining different ideas (Level 3) (Jansen et al., 2019).

3.2.3. The influence of context on students' meta-modelling knowledge

When determining students' expressed meta-modelling knowledge, it is important to be aware of the effect of the context in which a model is presented on the performance of the student. Krell, Upmeier zu Belzen and Krüger (2014) looked at the difference in students' understanding of models in a contextualized and a decontextualized context and found that students often lack an informed understanding of a model when a decontextualized context is used. They argue that the expressed level of understanding depends on both the aspect of models and modelling (*nature of models, purpose of models, etc.)* and the model itself. Krell and Krüger (2017) add that the subject is also an important context when assessing students' expressed meta-modelling knowledge; students show a more advanced understanding of models in a physics and chemistry context, and a less advanced understanding of models when they refer to a biology context. This might be caused by a difference in the way models are treated in class in these subjects (Gobert et al., 2011; Krell, Reinisch, & Krüger, 2015).

Considering the context of the model, Al-Balushi (2011) showed that the level of abstractness of a model relates to how students experience the model. Entities which can be captured on a photograph or a micrograph, such as an animal, or a red blood cell, feel more "real" to students than theoretical constructs such as electron clouds or photons. However, the level of abstractness can sometimes be

Table 3.1. Framework to assess students' meta-modelling knowledge of biological concept-process models. The left column shows the five aspects that are important when reasoning with biological models. For each of these aspects up to four levels of understanding have been defined, ranging from an initial level of understanding to an expert level of understanding. Categories in bold have specifically been added to assess students' meta-modelling knowledge of biological concept-process models (Jansen et al., 2019).

	Initial level	Level 1	Level 2	Level 3
Nature of models		- Model as copy - Model with great similarity - Model represents a (non-) subjective conception of the original **- Displays a process, its components and how they are related**	- Parts of the model are a copy - Model as a possible variant - Model as focused representation	- Model as hypothetical representation
Purpose of models		- Model for showing the facts **- Model for showing events**	- Model to identify relationships	- Model to examine abstract/ concrete ideas
Multiple models	- All models are the same - Various models of different originals - Only one final and correct model	- Different model object properties	- Focus on different aspects	- Different assumptions
Testing models	- No testing of models **- Perceiving schoolbooks or their authors as authorities providing absolute truth**	- Testing of material - Testing of basic requirements	- Comparison between original and model - Comparison and matching of original and model	- Testing hypotheses - Testing of hypotheses with research designs
Changing models	- No reason for alterations - Alteration of how different originals are represented	- Alterations to improve the model object - Alterations when there are errors in the model object - Alterations when basic requirements are not met	- Alterations when model does not match the original - Alterations due to new findings about the original	- Alterations due to findings from model experiments **- Alterations when the focus of the model shifts to a different aspect of the process**

misjudged by students. Krell et al. (2012) show that students often believe that a biomembrane or the water cycle for example can directly be observed, and that models of these concepts must be correct replications of the original. A study by Jansen et al., (2019) also suggests a difference in students' experience that can be related to the level of abstractness of the model. This study indicated a difference in students' meta-modelling knowledge for models on the cellular organizational level and models on the ecological organizational level, where students show a more advanced meta-modelling understanding for models on the cellular organizational level. These findings suggest that the biological level of organization and level of abstractness on which a process has been modelled has an influence on students' meta-modelling knowledge. The studies by Jansen et al. (2019) and Al-Balushi (2011) can be interpreted in the light of the classification by Tsui and Treagust (2013), who suggest a classification of the models on different levels of organization in biology, indicating a difference in students' expressed meta-modelling knowledge between macroscopic models (ecological level; visible to the naked eye) and (sub) microscopic models (cellular level; invisible to the naked eye).

Krell (2019) proposes that providing a modelling-purpose in the context of a model is essential to finding out whether students are capable of a Level 3 meta-modelling understanding. Since the same model can serve various purposes, both communicative and explanatory (Odenbaugh, 2005), it would be unrealistic to expect students to express their highest capable level of understanding in every given context. The descriptions of the aspect of purpose of models by Grosslight et al. (1991) and Crawford and Cullin (2005), as well as the aspects of use and prediction of models as described by Justi and Gilbert (2003), are combined in the framework as described by Upmeier zu Belzen et al. (2019). Three purposes are differentiated: models can serve to show facts about the original (Level 1), to describe and explain a known relationship in the original (Level 2) and as an instrument to predict the behaviour of an original (Level 3). Considering concept-processes this means that models can be used to visualize a process (Level 1), connect a certain process to other aspects or processes (Level 2) or formulate and test hypotheses (Level 3) (Jansen et al., 2019).

Krell (2019) assessed students' understanding for the aspect *testing models* and *changing models* and used three different modelling-purposes, corresponding to the three levels as shown in Table 3.1 for the aspect *purpose of models*: an aesthetic purpose (corresponding to Level 1), an explanatory purpose (corresponding to Level 2) and a research tool purpose (corresponding to Level 3). According to the results of this study, students tend to agree more often with a Level 3 statement when this statement is presented in a research tool purpose context.

3.2.4. Aim of the study

This study aims to describe eleventh grade pre-university students' meta-modelling knowledge considering biological concept-process models. Our study follows up on the research done by Krell (2019), who focused on the effect of the presence of a modelling purpose on students expressed meta-modelling knowledge considering the aspects *testing models* and *changing models*. The current study focuses on students' expressed meta-modelling knowledge considering the aspects *nature of models* and *multiple models,* while taking the context of the model - the presence of a modelling purpose and the difference between (sub) microscopic and macroscopic models - into account. This leads to the following research questions:

> *3.1: How does students' expressed meta-modelling knowledge, related to the aspects nature of models and multiple models, depend on the presence of different kinds of explicit modelling-purposes?*
>
> *3.2: How does students' expressed meta-modelling knowledge, related to the aspects nature of models and multiple models, differ between contexts showing macroscopic and (sub)microscopic concept-process models?*

3.3. Method

A task-based test was developed to describe eleventh grade pre-university students' expressed meta-modelling knowledge considering biological concept-process models. Two repeated measure ANOVAs were performed to find out whether the type of model ((sub)microscopic or microscopic) and the presence of an explicit modelling purpose influenced students' expressed meta-modelling knowledge.

3.3.1. Participants

Participants were 430 Dutch eleventh grade pre-university level students with a major in biology (16-18 years old) from sixteen schools in different areas in The Netherlands. Forty-three students (10%) did not complete the entire test and were removed from the analysis, resulting in 387 complete tests included in this study.

Participating schools were recruited via an online announcement. For the areas in The Netherlands that were not yet represented, the researchers contacted schools personally. The test was filled out in class under the teachers' supervision. Informed consent was obtained from all participating students.

3.3.2. Instrument

An online test to assess students' meta-modelling knowledge considering the aspects *nature of models* and *multiple models* was developed. The test was based on the design described by Krell (2019), who also assessed students' meta-modelling knowledge, but focused on the aspects *testing models* and *changing models* (Table 3.1). The tasks in the design as described by Krell (2019) are different from forced-choice tasks that were used in previous research on this topic (Gogolin & Krüger, 2018; Krell et al., 2012). In a forced-choice task in which students need to choose between statements, they can only display (or choose) one level of understanding. Thus, the result of these forced-choice tasks only shows the preferred level of understanding. Since we are interested in finding out students' expressed meta-modelling knowledge for all three levels of understanding, we deviated from the forced-choice task and followed the design by Krell (2019) to develop a closed-question test for the assessment of students' expressed meta-modelling knowledge on all three levels of understanding. The considerations as described about the context of the model (level of biological organization and the presence or absence of different modelling purposes), the aspects of focus (*nature of models* and *multiple models*) and the corresponding levels as described in Table 3.1 are integrated in all tasks in this test and are referred to as 'dimensions' within the tasks. The dimensions have been specified as follows: *1. Type of model, 2. Modelling purpose, 3. Aspect of meta-modelling knowledge and 4. Level of understanding*.

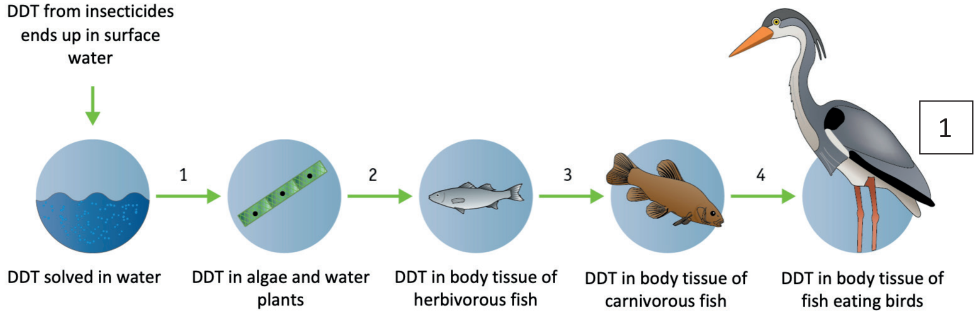
For dimension 1, Type of model, six different biological concept-process models were used in the test: three microscopic models and three macroscopic models. Considering dimension 2, Modelling purpose, students were first faced with the six different models without an accompanying modelling purpose. In the second part of the test, students reflected on the same six models, but this time in combination with an explicit modelling purpose. In this second part of the test, the given purpose was either aesthetic (Level 1), representative (Level 2) or research minded (Level 3). Relating to dimension 3 and 4, Aspect of meta-modelling knowledge and Level of understanding, students were asked to judge whether given statements are appropriate in the context (type of model and - if present - the given purpose) that was presented to them (yes/no questions). These statements were either based on the meta-modelling aspect *nature of models,* or *multiple models*. Three statements per context were formulated, each based on the three levels of understanding as described in Table 3.1. Hence, students' judgement about these statements was interpreted as their expressed meta-modelling knowledge. In separate tasks, each of the concept-process models that was used in the test was combined with four different types of modelling purpose (no given modelling purpose, aesthetic purpose, representative purpose and research purpose), two aspects of meta-modelling knowledge (*nature of models* and *multiple models*) and three levels of understanding (Level 1, Level 2 and Level 3). The systematic combination of the dimensions resulted in 6 x 4 x 2 x 3 = 144 possible statements in total. Table 3.2 shows the set-up of the test with possible combinations of dimensions relating to one of the concept-process models that was used in this study. All tasks were developed based on the same abstract template, providing elements containing the four different dimensions. Furthermore, all tasks were critically discussed and optimized by science education researchers with knowledge on models, modelling and assessment in science education. Figure 3.2 shows an example of one of the tasks in the test. A link to the full online test can be found in Appendix C.

Table 3.2. Test design for the development of the test, where the resulting types of statements for one of the models are shown. Each column represents one of the four dimensions that were taken into account: type of model, purpose of modelling, aspect of meta-modelling knowledge and level of understanding. In *italic* an example of one task consisting of three statements is visualized. The resulting task is shown in Figure 3.2.

Type of model	Modelling purpose	Aspect of meta-modelling knowledge	Level of understanding
Bioaccumulation (macroscopic model)	No given modelling-purpose	Nature of models	1, 2, 3
		Multiple models	1, 2, 3
	Aesthetic	Nature of models	1, 2, 3
		Multiple models	1, 2, 3
	Representative	Nature of models	1, 2, 3
		Multiple models	1, 2, 3
	Research	Nature of models	1, 2, 3
		Multiple models	*1, 2, 3*

Example item

Frank is doing research on water contamination. He discovered that in areas where the water is contaminated with DDT (a chemical insecticide), many fish eating birds are dying. He creates the following model.



Maud is also doing research on water contamination with DDT. 2

She creates a model of the same process. When can the model that Maud creates be defined as a different model than Franks' model? 3

Answer for each of the statements if you agree (yes) or disagree (no) 4

When Maud creates a model for an area that is contaminated with DDT, but where different species of fish are living.
[ ] Yes
[ ] No

When Maud creates a model where the focus is more on step 2 of the process.
[ ] Yes
[ ] No

When Maud has different thoughts on how DDT ends up in fish eating birds and creates a model reflecting these thoughts.
[ ] Yes
[ ] No

Figure 3.2. An example of one of the closed question tasks in the test, showing a model of bioaccumulation. The four dimensions are highlighted and numbered (not visible in the student version of the test). In this example the four dimensions are integrated as follows: Type of model (1): macroscopic, showing the process of bioaccumulation. Modelling purpose (2): research purpose, stating that Maud is doing research on water contamination. Aspect of meta-modelling knowledge (3): multiple models, since Maud is creating a different model about the same topic. Level of understanding (4): each statement represents one of the levels of understanding in the order: Level 1, Level 2 and Level 3 (the order of the levels differs between different tasks). The model present in this task is reprinted and translated with permission from Bos et al. (2012).

Participants were presented with 12 closed question tasks (see Table 3.3 for an example of a set-up of a complete test). Half of these tasks included a macroscopic model, the other half a (sub)microscopic model. Furthermore, the first six tasks were presented without an explicit modelling-purpose, the second six tasks did contain a modelling-purpose: two tasks presented an aesthetic (Level 1) modelling-purpose, two tasks a representative (Level 2) modelling-purpose and two tasks a research (Level 3) modelling-purpose. Since we used six different concept-process models, participants encountered each model twice: once with a non-explicit modelling-purpose (no purpose-set) and once with an explicit modelling purpose (explicit purpose-set). In order to prevent an influence of explicit modelling-purposes on the non-explicit modelling purpose tasks, all tasks in the no purpose-set were presented before all tasks in the explicit purpose-set. Within each set, the order of the tasks was randomized to prevent order effects.

3

Table 3.3. An example of a set-up of a complete test for one participant. The test consisted of two sets: in the no purpose-set students were presented with models without a modelling purpose. In the explicit purpose-set students were presented with the same models as in the no purpose-set, but combined with a modelling purpose (aesthetic, representative, or research purpose). Half of the models were macroscopic, the other half (sub)microscopic. Each task within a set contained three statements, referring to either Level 1, 2 or 3 for the aspects *nature of models* or *multiple models*. Within each set the order of the tasks was randomized to prevent order effects.

Set	Task	Model	Modelling-purpose	Aspect of meta-modelling knowledge	Present statements
No purpose-set	1	I: macroscopic	none	nature of models	Level 1, 2 and 3
	2	II: macroscopic	none	multiple models	Level 1, 2 and 3
	3	III: macroscopic	none	nature of models	Level 1, 2 and 3
	4	IV: (sub)microscopic	none	multiple models	Level 1, 2 and 3
	5	V: (sub)microscopic	none	nature of models	Level 1, 2 and 3
	6	VI: (sub)microscopic	none	multiple models	Level 1, 2 and 3
Explicit purpose-set	7	I: macroscopic	aesthetic	multiple models	Level 1, 2 and 3
	8	II: macroscopic	representative	nature of models	Level 1, 2 and 3
	9	III: macroscopic	research tool	multiple models	Level 1, 2 and 3
	10	IV: (sub)microscopic	aesthetic	nature of models	Level 1, 2 and 3
	11	V: (sub)microscopic	representative	multiple models	Level 1, 2 and 3
	12	VI: (sub)microscopic	research tool	nature of models	Level 1, 2 and 3

3.3.3. Pilot

A pilot with nine pre-university education students (six female, three male) was conducted to assess the validity of the test. The participants were asked to complete the test individually. After they completed the test, a 30-minute semi-structured interview was conducted. The goal of the interview was four-fold: 1) to find out whether the tasks were clear and well understood by the participants 2) to follow the participants' way of reasoning while selecting the answers 3) to observe whether any previous tasks influenced the participants' answers on following tasks 4) to find out whether the length and the number of questions was adequate.

The outcome of the pilot resulted in minor linguistic changes; the use of personal names instead of function titles (e.g., *Lisa* instead of *the researcher*); and a change in the position of the modelling purpose in the task (below the model instead of above). Also, we decided to make the fact that students would encounter each model twice, but in a different context, more explicit. Therefore, a screen was added to the online test in between the two sets of tasks, mentioning that the models would be repeated and that the students should read the text carefully since the models would be presented in a different context.

3.3.4. Test administration

All participants received a short introduction about the procedure of the experiment by the researcher (either in real-life or with a short video-clip). Then, the test was administered to the participants digitally and carried out individually. In order for students to be able to see every model in detail, only devices with large screens, such as laptops or computers, were used.

3.3.5. Data management

As in the study done by Krell (2019), students' agreement with statements was used as a measure for their expressed meta-modelling knowledge. In order to compare students' agreement to statements in the no purpose-set with their agreement to statements in the explicit purpose-set, our data had to be normalized. Scores for each statement were labelled 1 for agree/yes and 0 for disagree/no. In the no purpose-set within the test, only the modelling purpose 'no purpose' was represented (see Table 3.3). This set within the test contained three statements per combination of variables (e.g., no purpose + *nature of models* + Level 1). The total agreement score per combination of variables per student for this set within the test can therefore range from 0 – 3, where a student reaches a total agreement score of 0 for a combination of variables when agreeing with none of the three statements, and a total agreement score of 3 when agreeing with all three statements. Scores were normalized to fall between 0 and 1.

The explicit purpose-set within the test contained three different modelling purposes (aesthetic, explanatory and research tool) (see Table 3.3). This part of the test contained one statement per combination of variables (e.g., explanatory purpose + *multiple models* + Level 2). The total agreement score per combination of variables per student for the explicit purpose-set within the test is either 0 or 1, since the student either disagreed or agreed with the statement.

In order to compare students' agreement to statements in contexts containing a macroscopic model with their agreement to statements in contexts containing a (sub)microscopic model, our data again had to be normalized. Since we wanted to exclude the possible effect of different types of explicit modelling purposes on our results, only data from the 'no purpose-set' was used for this analysis. The no-purpose set within the test contained three statements per combination of variables (e.g., macroscopic model + *nature of models* + Level 1). Due to the randomization procedure of the tasks within both sets of the test, students were presented with either one or two statements per combination of variables. The total agreement score per combination of variables per student can therefore range from 0 – 2. Scores were normalized to fall between 0 and 1, while taking the number of answered statements per combination of variables into account.

To determine the presence of 'unnatural patterns' (e.g., only yes/no for all statements) and preclude the possibility of random answers, all possible answer patters were defined and individual student answers were matched to these patterns. An answer pattern is the agreement pattern for the three statements of one task (Figure 3.2). If a student agrees with the statement on Level 1 and Level 2, but not with the statement on Level 3, the answer pattern is agree-agree-disagree. Eight different answer patterns are possible (Level 1 agree/disagree, Level 2 agree/disagree, Level 3 agree/disagree, thus 2 x 2 x 2 = 8 patterns).

3.3.6. Data analysis

To answer RQ1, we performed a repeated measures ANOVA to determine the effect of different kinds of explicit modelling-purposes on the agreement score for statements on the three levels of meta-modelling knowledge, considering the aspects *nature of models* and *multiple models* (software used: IBM SPSS Statistics, version 25). In the repeated measures ANOVA 'no purpose' was labelled as a category of modelling purposes, next to the three explicit modelling purposes (aesthetic, explanatory and research tool). In this repeated measures ANOVA we treated the three dimensions Modelling purpose (no purpose, aesthetic purpose, explanatory purpose and research tool purpose), Aspect of meta-modelling knowledge (*nature of models* and *multiple models*) and Level of understanding (Level 1, 2 and 3) as independent variables, and the students' agreement score as the dependent variable. Post-hoc pairwise comparisons (Bonferroni) were carried out for each of the independent variables to see whether students' agreement scores significantly differed within these variables.

Mauchly's test indicated that the assumption of sphericity had been violated for the variables Level of understanding $\chi^2(2) = 82.43$, $p < .001$ and Modelling purpose $\chi^2(5) = 39.44$, $p < .001$. Therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .84$ for Level of understanding and $\varepsilon = .94$ for Modelling purpose).

Considering the combination of specific independent variables, Mauchly's test indicated that the assumption of sphericity had been violated for the interactions Level of understanding * Aspect of focus $\chi^2(2) = 40.79$, $p < .001$, Level of understanding * Modelling Purpose $\chi^2(20) = 118.34$, $p < .001$ and Aspect of meta-modelling knowledge * Modelling purpose $\chi^2(5) = 54.40$, $p < .001$. Therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .91$ for Level of understanding * Aspect of meta-modelling knowledge, $\varepsilon = .91$ for Level of understanding * Modelling purpose and $\varepsilon = .93$ for Aspect of meta-modelling knowledge * Modelling purpose).

Considering the combination of all independent variables, Mauchly's test indicated that the assumption of sphericity had been violated for the interaction Aspect of meta-modelling knowledge * Level of understanding * Modelling purpose $\chi^2(20) = 183.51$, $p < .001$. Therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .87$).

To answer RQ2 we performed a second repeated measures ANOVA to determine the effect of macroscopic and (sub)microscopic models on the agreement score for statements on the three levels of meta-modelling knowledge, considering the aspects *nature of models* and *multiple models* (software used: IBM SPSS Statistics, version 25). In order to exclude the possible effect of different types of explicit modelling purposes on our result, we only used data from the 'no purpose-set' within the test for this second repeated measures ANOVA. In this repeated measures ANOVA we treated the three dimensions Type of model (macroscopic and (sub) microscopic), Aspect of meta-modelling knowledge (*nature of models* and *multiple models*) and Level of understanding (Level 1, 2 and 3) as independent variables, and the students' agreement score as the dependent variable. Mauchly's test indicated that sphericity could be assumed for the variable Level of understanding.

To find out whether specific combinations of the independent variables (Level of understanding, Aspect of meta-modelling knowledge and Type of model) have an effect on students' agreement scores, interaction effects for each of the independent variables were studied. Mauchly's test indicated that the assumption of sphericity had been violated for the interactions Level of understanding * Aspect of meta-modelling knowledge $\chi^2(2) = 25.16$, $p < .001$ and Level of understanding * Type of model $\chi^2(2) = 8.02$, $p = .018$. Therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .94$ for Level of understanding * Aspect of meta-modelling knowledge and $\varepsilon = .98$ for Level of meta-modelling knowledge * Type of model).

Considering the combination of all independent variables, Mauchly's test indicated that the assumption of sphericity had been violated for the interaction Level of understanding * Aspect of meta-modelling knowledge * Type of model $\chi^2(2) = 4.82$, $p = .090$. Therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .99$).

3.4. Results

The results will be described in line with the two research questions.

> *3.1: How does students' expressed meta-modelling knowledge, related to the aspects nature of models and multiple models, depend on the presence of different kinds of explicit modelling-purposes?*

Table 3.4. Mean students' agreement scores for the four purposes (no purpose, aesthetic purpose, representative purpose and research purpose) and three levels (Level 1, 2 and 3) considering the aspect *nature of models* and *multiple models,* where MAS = mean agreement score and SD = standard deviation, (n=387).

	Nature of models				Multiple models				Overall						
	No purpose MAS (SD)	Aesthetic purpose MAS (SD)	Representative purpose MAS (SD)	Research tool purpose MAS (SD)	No purpose MAS (SD)	Aesthetic purpose MAS (SD)	Representative purpose MAS (SD)	Research tool purpose MAS (SD)	No purpose MAS (SD)	Aesthetic purpose MAS (SD)	Representative purpose MAS (SD)	Research tool purpose MAS (SD)	All purposes MAS (SD) MAS (SD)	*Nature of models* MAS (SD)	*Multiple models* MAS (SD)
Level 1	0.76 (0.25)	0.80 (0.40)	0.72 (0.45)	0.75 (0.43)	0.19 (0.22)	0.22 (0.42)	0.26 (0.44)	0.26 (0.44)	0.43 (0.24)	0.51 (0.41)	0.49 (0.45)	0.50 (0.38)	0.50 (0.38)	0.76 (0.38)	0.23 (0.38)
Level 2	0.80 (0.24)	0.78 (0.41)	0.81 (0.39)	0.79 (0.41)	0.53 (0.38)	0.53 (0.50)	0.55 (0.49)	0.57 (0.50)	0.67 (0.31)	0.66 (0.46)	0.68 (0.44)	0.68 (0.46)	0.67 (0.42)	0.80 (0.36)	0.55 (0.47)
Level 3	0.60 (0.33)	0.63 (0.48)	0.65 (0.48)	0.66 (0.48)	0.84 (0.23)	0.87 (0.34)	0.84 (0.37)	0.88 (0.33)	0.72 (0.28)	0.75 (0.41)	0.75 (0.43)	0.77 (0.41)	0.75 (0.38)	0.64 (0.44)	0.86 (0.32)
Overall	0.72 (0.27)	0.74 (0.43)	0.73 (0.44)	0.74 (0.44)	0.52 (0.28)	0.54 (0.42)	0.55 (0.43)	0.57 (0.42)	0.62 (0.28)	0.64 (0.43)	0.64 (0.44)	0.66 (0.43)	0.64 (0.39)	0.73 (0.40)	0.55 (0.39)

A repeated measures ANOVA was used to look into the effect of all different categories of modelling purposes (no purpose, aesthetic, explanatory and research tool) on students' agreement scores for statements on the three levels of meta-modelling knowledge (Level 1, 2 and 3), considering the aspects nature of models and multiple models. Table 3.4 shows the mean agreement scores considering each of the modelling purposes for the aspects nature of models and multiple models.

The results show that all independent variables have a significant influence on the students' agreement score, with $F(1.68, 647.25) = 297.29$, $p < .001$ for Level of understanding, $F(1, 386) = 424.02$, $p < .001$ for Aspect of meta-modelling knowledge, and $F(2.83, 1093.66) = 3.25$, $p = .023$ for Modelling purpose.

Post-hoc pairwise comparisons (Bonferroni) for the independent variable Level of understanding show an increase in agreement score from Level 1 to Level 3, where agreement scores are significantly higher for Level 2 than for Level 1 ($DM = .294$, $p < .001$) and significantly higher for Level 3 than for Level 2 ($DM = .219$, $p < .001$).

Pairwise comparisons for the independent variable Aspect of meta-modelling knowledge show a significantly higher agreement score for the aspect *nature of models* than for the aspect *multiple models* ($DM = .251$, $p < .001$). Pairwise comparisons for the independent variable Modelling purpose only show a significant difference in agreement score between statements in tasks containing 'no purpose' and tasks containing a 'research tool purpose' ($DM = .030$, $p < .001$), where students' agreement score is higher for statements in tasks containing a 'research tool purpose' than for statements in tasks containing 'no purpose'.

To find out whether specific combinations of the independent variables have an effect on students' agreement scores, interaction effects for each of the independent variables were studied. Results show a significant interaction effect for the variables Level of understanding and Aspect of meta-modelling knowledge $F(1.82, 701.48) = 260.38$ $p < .001$. As can be seen in Figure 3.3, the agreement scores for the two aspects show a different pattern, where both the difference in agreement scores between Level 1 and Level 2 $F(1, 386) = 547.92$, $p < .001$, and Level 1 and Level 3 $F(1, 386) = 327.61$, $p < .001$ differ between the two aspects.

Results also show a significant interaction effect for the variables Level of understanding and Modelling purpose $F(5.49, 2117.63) = 2.42$, $p = .029$. Figure 3.4 shows this interaction, where only the difference in agreement scores between

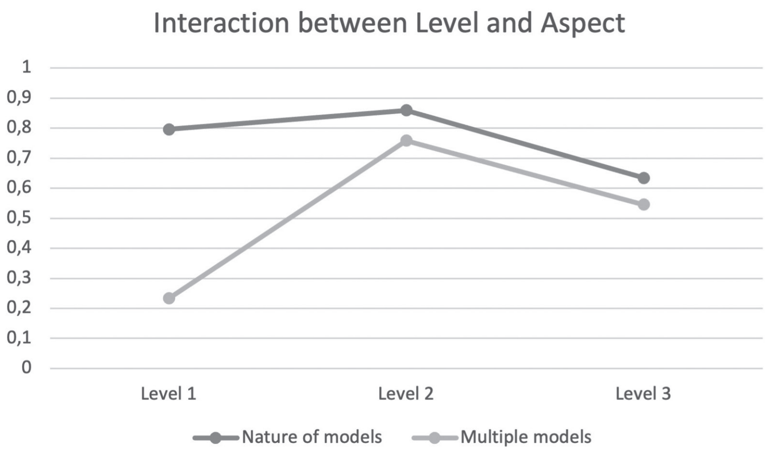


Figure 3.3. The interaction considering students' agreement scores between the variables Level of understanding (level 1, 2 or 3) and Aspect of meta-modelling knowledge (*multiple models* or *nature of models*). The average agreement score ranges from 0 (no agreement) to 1 (agreement with all statements). The difference in agreement scores between Level 1 and Level 2 $F(1, 386) = 547.92$, $p < .001$, and Level 1 and Level 3 $F(1, 386) = 327.61$, $p < .001$ differs between the two aspects.

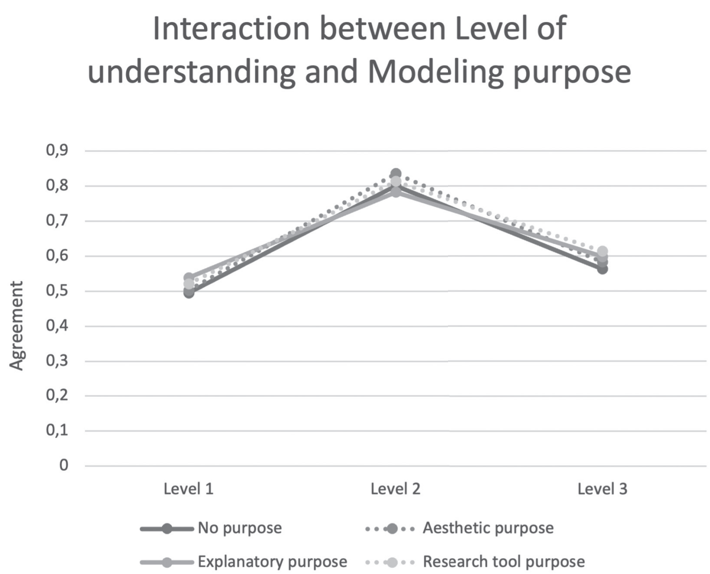


Figure 3.4. The interaction considering students' agreement scores between the variables Level of understanding (level 1, 2 or 3) and Modelling purpose (no purpose, aesthetic purpose, representative purpose and research tool purpose). The average agreement scores ranges from 0,4 (partial agreement) to 1 (agreement with all statements). The difference in agreement scores between contexts with no modelling purpose and a representative purpose is significant when comparing students' answers for Level 1 and Level 2 statements $F(1, 386) = 6.74$, $p = .010$ (uninterrupted lines).

contexts with no modelling purpose and an explanatory purpose is significant when comparing students' answers for Level 1 and Level 2 statements $F(1, 386) = 6.74$, $p = .010$.

To find out whether the combination of all independent variables has an effect on students' agreement scores, the interaction effect for this combination of variables was studied. Results show no significant interaction of the three independent variables (Aspect of meta-modelling knowledge, Level of understanding and Modelling purpose) on the agreement score $F(5.21, 2011.73) = 1.274$, $p = .27$. Table 3.5 summarizes the results from the repeated measures ANOVA.

3.2: How does students' expressed meta-modelling knowledge, related to the aspects nature of models and multiple models, differ between contexts showing macroscopic and (sub)microscopic concept-process models?

A repeated measures ANOVA on data from the no-purpose set within the test was used to look into the effect of contexts containing macroscopic or (sub)microscopic models on students' agreement scores for statements on the three levels of meta-modelling knowledge (Level 1, 2 and 3), considering the aspects *nature of models* and *multiple models*. Table 3.6 shows the mean agreement scores considering both types of models for the aspects *nature of models* and *multiple models*.

We first focused on the effect of each of the independent variables (Level of understanding, Aspect of meta-modelling knowledge and Type of model) on students' agreement scores. Since this second repeated measures ANOVA was run on data from only the no purpose-set within the test, results for all combinations of variables considering this dataset are considered.

Table 3.5. Results of the repeated measures ANOVA relating to RQ1 with Level of understanding, Aspect of meta-modelling knowledge and Modelling purpose as independent variables and the students' agreement score as dependent variable. *SS* = sum of squares, *df* = degrees of freedom and *MS* = mean square.

		SS	***df***	***MS***	***F***	***p***	**partial h^2**
Aspect	Sphericity Assumed	146461	1	146461	424015	.000	.523
Error (Aspect)	Sphericity Assumed	133330	386	.345			
Level	Greenhouse-Geisser	144200	1677	85995	297293	.000	.435
Error (Level)	Greenhouse-Geisser	187226	647256	.289			
Purpose	Greenhouse-Geisser	1123	2833	.396	3251	.023	.008
Error (Purpose)	Greenhouse-Geisser	133308	1093657	.122			
Aspect x Level	Greenhouse-Geisser	112724	1817	62028	260378	.000	.403
Error (Aspect x Level)	Greenhouse-Geisser	167109	701483	.238			
Aspect x Purpose	Greenhouse-Geisser	.088	2776	.032	.237	.856	.001
Error (Aspect x Purpose)	Greenhouse-Geisser	144287	1071572	.135			
Level x Purpose	Greenhouse-Geisser	2014	5486	.367	2417	.029	.006
Error (Level x Purpose)	Greenhouse-Geisser	321597	2117633	.152			
Aspect x Level x Purpose	Greenhouse-Geisser	1014	5212	.195	1274	.271	.003
Error (Aspect x Level x Purpose)	Greenhouse-Geisser	307153	2011725	.153			

Table 3.6. Mean students' agreement scores for the two types of models (macroscopic and (sub)microscopic) and three levels (Level 1, 2 and 3) considering the aspects *nature of models* and *multiple models,* where MAS = mean agreement score and SD = standard deviation, (n=387).

	Nature of models		**Multiple models**		**Overall**				
	(Sub)microscopic MAS (SD)	Macroscopic MAS (SD)	(Sub)microscopic MAS (SD)	Macroscopic MAS (SD)	(Sub) microscopic MAS (SD)	Macroscopic MAS (SD)	(Sub) microscopic + macroscopic MAS (SD)	*Nature of models* MAS (SD)	*Multiple models* MAS (SD)
Level 1	0.64 (0.42)	0.89 (0.26)	0.24 (0.36)	0.11 (0.24)	0.45 (0.39)	0.69 (0.29)	0.48 (0.32)	0.77 (0.34)	0.18 (0.30)
Level 2	0.79 (0.37)	0.80 (0.35)	0.53 (0.45)	0.53 (0.46)	0.65 (0.41)	0.73 (0.39)	0.66 (0.41)	0.79 (0.37)	0.53 (0.46)
Level 3	0.66 (0.44)	0.61 (0.42)	0.83 (0.32)	0.89 (0.27)	0.52 (0.38)	0.67 (0.39)	0.74 (0.36)	0.63 (0.43)	0.86 (0.30)
Overall	0.70 (0.41)	0.77 (0.34)	0.53 (0.38)	0.51 (0.32)	0.62 (0.40)	0.64 (0.33)	0.63 (0.36)	0.74 (0.38)	0.52 (0.35)

In line with the results from our first repeated measures ANOVA, the results from this second repeated measures ANOVA show that the variables Level $F(2,772) = 171.22$, $p < .001$ and Aspect $F(1,386) = 377.25$, $p < .001$ have a significant influence on students' agreement score. The results of this second repeated measures ANOVA also show that the independent variable Type of model has a significant influence on students' agreement scores $F(1,386) = 6.40$, $p = .012$, where the agreement score for macroscopic models is higher than for (sub)microscopic models ($DM = .023$, $p = .012$).

Interaction effects for each of the independent variables were studied to find out whether specific combinations of the independent variables (Level of understanding, Aspect of meta-modelling knowledge and Type of model) have an effect on students' agreement scores. In line with the results from our first repeated measures ANOVA, results for this second repeated measures ANOVA showed a significant interaction for the variables Level of understanding * Aspect of meta-modelling knowledge $F(1.88,726.07) = 373.16$, $p < .001$. Results also showed a significant interaction for the variables Type of model * Aspect of meta-modelling knowledge $F(1,386) = 28.97$, $p < .001$. Figure 3.5 shows that students more often agree with statements related to the aspect *multiple models* when confronted with a context containing a (sub)microscopic model than when confronted with a context containing a macroscopic model. The opposite response pattern in students' agreement is found when looking at statements related to the aspect *nature of models*.

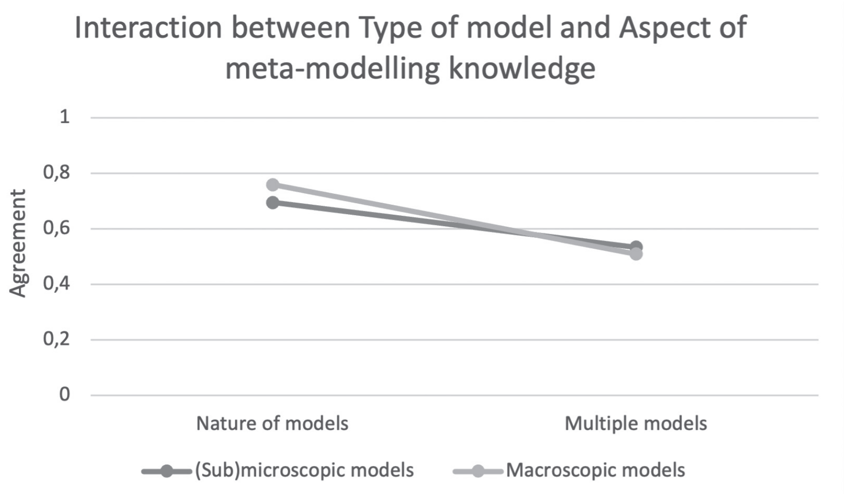


Figure 3.5. The interaction considering students' agreement scores between the variables Type of model ((sub)microscopic or macroscopic) and Aspect of meta-modelling knowledge (*multiple models* or *nature of models*). The average agreement score ranges van 0 (no agreement) to 1 (agreement with all statements) $F(1,386) = 28.97$, $p < .001$.

Our results showed no significant interaction for the variables Type of model * Level of understanding $F(1.96,756.42) = 2.82$, $p = .061$.

To find out whether the combination of all independent variables (Level of understanding, Aspect of meta-modelling knowledge and Type of model) has an effect on students' agreement scores, the interaction effect for this combination of variables was studied. Results show a significant interaction of the three independent variables on the agreement score $F(1.98, 762.51) = 59.03$, $p < .001$. A significant effect is found when comparing agreement scores for statements on Level 1 and Level 3 $F(1, 386) = 106.18$, $p < .001$, Level 1 and Level 2 $F(1, 386) =$

55.08, $p < .001$ and Level 2 and Level 3 $F(1, 386) = 5.62$, $p = .018$, while comparing the agreement scores for statements relating to the aspects *nature of models* and *multiple models,* and for contexts containing macroscopic and (sub)microscopic models. Figures 3.6 and 3.7 visualize this interaction, where Figure 3.6 shows the interaction between the variables Type of model and Level of understanding for the aspect *nature of models,* and Figure 3.7 shows this interaction for the aspect *multiple models*. For the aspect *nature of models* students' agreement scores show a difference in pattern considering (sub)microscopic models and macroscopic models. Students' agreement scores for (sub)microscopic models are highest for Level 1 statements and lowest for Level 3 statements, while students' agreement scores for macroscopic models are highest for Level 2 statements and lower for Level 1 and Level 3 statements. For the aspect *multiple models* students' agreement scores considering (sub)microscopic models and macroscopic models both show a similar pattern, but students' agreement scores for Level 1 statements are highest for (sub)microscopic models, while students' agreement scores for Level 3 statements are highest for macroscopic models. Table 3.7 summarizes the results from the repeated measures ANOVA.

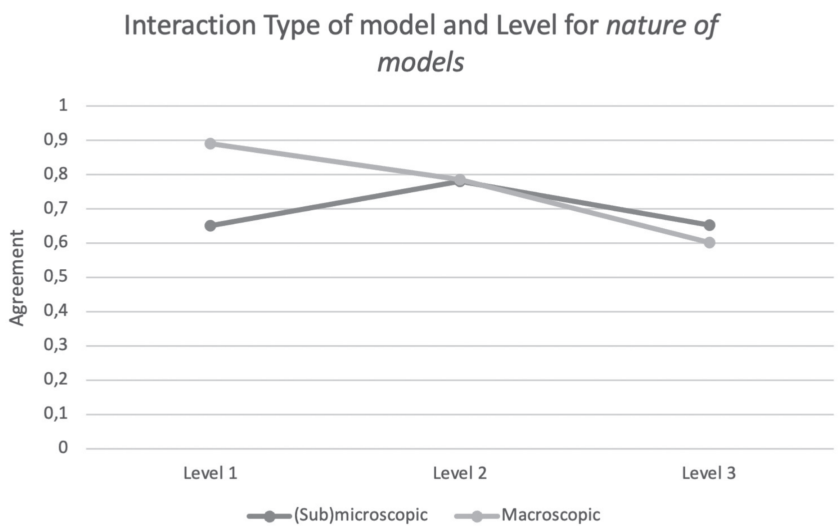


Figure 3.6. The interaction considering students' agreement scores between the variables Type of model ((sub)microscopic or macroscopic) and Level of understanding (Level 1, 2 or 3) for the aspect *nature of models*.

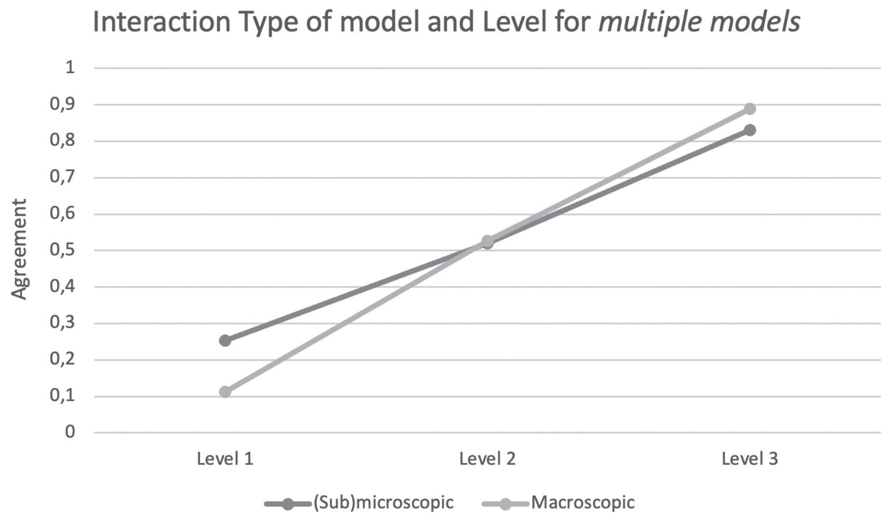


Figure 3.7. The interaction considering students' agreement scores between the variables Type of model ((sub)microscopic or macroscopic) and Level of understanding (Level 1, 2 or 3) for the aspect *multiple models*.

To determine the presence of 'unnatural patterns' in student answers, eight possible answer patterns were defined. Table 3.8 and Table 3.9 show the possible answer patterns (A - F), categorized per modelling aspect and given modelling purpose, and the percentage of students that followed each answer pattern.

For the modelling-aspect *nature of models*, there are two most common patterns. The first most common pattern (34 – 39%) is pattern H, which means agreeing with all statements. The second most common pattern (21 – 27% of all answers) is pattern F, which means agreeing with Level 1 and Level 2 statements, while disagreeing with Level 3 statements. The least common pattern is pattern A, meaning disagreement with all statements (1 – 2 % of all answers).

For the modelling-aspect *multiple models*, the two most common patterns are pattern C (34 – 38%), which means disagreeing with Level 1 statements and agreeing with Level 2 and Level 3 statements, and pattern D (28 – 33%), which means disagreeing with Level 1 and Level 2 statements and agreeing with Level 3 statements. The least common pattern (1 – 2%) is pattern E, meaning agreement with Level 1 statements and disagreement with Level 2 and Level 3 statements.

Table 3.7. Results of the repeated measures ANOVA relating to RQ2 with Level of understanding, Aspect of meta-modelling knowledge and Modelling purpose as independent variables and the students' agreement score as dependent variable. *SS* = sum of squares, *df* = degrees of freedom and *MS* = mean square.

		SS	***df***	***MS***	***F***	***p***	**partial h²**
Type of model	Sphericity Assumed	.605	1	.605	6.400	.012	.016
Error (Type of model)	Sphericity Assumed	36.478	386	.095			
Aspect	Sphericity Assumed	51.280	1	51.280	377.245	.000	.494
Error (Aspect)	Sphericity Assumed	52.470	386	.136			
Level	Sphericity Assumed	60.021	2	30.011	171.221	.000	.307
Error (Level)	Sphericity Assumed	135.312	772	.175			
Type of model x Aspect	Sphericity Assumed	2.653	1	2.653	28.973	.000	.070
Error (Type of model x Aspect)	Sphericity Assumed	35.347	386	.092			
Type of model x Level	Greenhouse-Geisser	.702	1.960	.358	2.822	.061	.007
Error (Type of model x Level)	Greenhouse-Geisser	95.965	756.415	.127			
Aspect x Level	Greenhouse-Geisser	130.505	1.881	69.380	373.159	.000	.492
Error (Aspect x Level)	Greenhouse-Geisser	134.995	726.070	.186			
Type of model x Aspect x Level	Greenhouse-Geisser	12.932	1.975	6.547	59.029	.000	.133
Error (Type of model x Aspect x Level)	Greenhouse-Geisser	84.568	762.507	.111			

Table 3.8. Answer patterns for the aspect *nature of models*. For each modelling purpose the percentage of students that answered according to each pattern is shown, where "+" indicates agreement and "-" indicates disagreement (*n = 387)*.

	Answer pattern			Modelling purpose			
	Level 1	Level 2	Level 3	No purpose (%)	Aesthetic purpose (%)	Explanatory purpose (%)	Research tool purpose (%)
A	-	-	-	2	3	4	2
B	-	+	-	9	6	7	6
C	-	+	+	35	36	34	38
D	-	-	+	33	32	28	28
E	+	-	-	2	1	1	1
F	+	+	-	2	2	4	3
G	+	-	+	10	10	12	11
H	+	+	+	6	9	9	11

Table 3.9. Answer patterns for the aspect *multiple models*. For each modelling purpose the percentage of students that answered according to each pattern is shown, where "+" indicates agreement and "-" indicates disagreement (*n = 387)*.

	Answer pattern			Modelling purpose			
	Level 1	Level 2	Level 3	No purpose (%)	Aesthetic purpose (%)	Explanatory purpose (%)	Research tool purpose (%)
A	-	-	-	1	1	2	1
B	-	+	-	8	4	7	8
C	-	+	+	12	11	15	11
D	-	-	+	3	5	4	5
E	+	-	-	6	6	5	3
F	+	+	-	26	27	21	23
G	+	-	+	10	11	8	13
H	+	+	+	34	36	39	36

3.5. Discussion

Understanding how models are created and used in science is an important aspect of scientific literacy. In order to foster students' meta-modelling knowledge, many model-based learning approaches have been developed (e.g., Gilbert & Justi, 2016; Krell et al., 2012; Quillin & Thomas, 2015; Schwarz & White, 2005; Treagust et al., 2002). To provide a basis for developing such model-based learning approaches that focus on *concept-process models* specifically, this study aims to describe eleventh grade pre-university level students' expressed meta-modelling knowledge considering biological concept-process models while taking the type of model and context of the model into account. The framework used in this study (Jansen et al., 2019) to develop a test for the assessment of students' expressed meta-modelling knowledge defines important aspects of meta-modelling knowledge for biological models. The descriptions of two of the aspects in this framework (*nature of models* and *multiple models*) and corresponding levels (Level 1, Level 2 and Level 3) were used in this study to develop a test. Students had the choice to agree or disagree with statements that were formulated in relation to these two aspects and three levels. As previously formulated by Krell (2019), students' agreement with statements can be interpreted as their expressed level of meta-modelling knowledge, since each statement is directly related to one of the aspects and levels as described in the framework. The interviews during our pilot study where students had to substantiate their choices also confirmed this relationship between students' agreement with statements and their expressed meta-modelling knowledge. For example, students explained that they agreed with a Level 2 statement for the aspect *multiple models,* because "things had been added or left out in order to focus on a part of the process" or agreed with a Level 3 statement for the aspect *nature of models* because "a hypothesis means that the model just shows the way the researcher thinks about the process, but they still need to test it". By using the framework as described by Jansen et al. (2019) to investigate students' current understanding of models, we hope to gain more insight in which aspects of model-based reasoning need to be fostered to enhance students' expressed meta-modelling knowledge and stimulate scientific literacy.

To find out which answer patterns occurred most often and to preclude the possibility that students randomly agreed or disagreed with statements, every individual student answer pattern was scored (Table 3.8 and Table 3.9). If students would have randomly agreed or disagreed with the given statements, we would expect an even distribution among the possible answer patterns. However, for both the aspects *nature of models* and *multiple models* two answer patterns stood out with more students following these patterns than other possible patterns, suggesting that students did not randomly agree or disagree with statements. Unnatural answer patterns with students either agreeing or disagreeing with all statements have also been studied. Even though only a small percentage (1 - 4%) of students disagreed with all statements in a task for both the aspects *nature of models* and *multiple models*, we did find a large percentage (34 – 39%) of students agreeing with all statements in a task for the aspect *nature of models*. However, there is only a small percentage (6 – 11%) of students showing this answer pattern (agreeing with all statements) for the aspect *multiple models.* Since tasks related to the aspects *nature of models* and *multiple models* were randomized in the test, this difference in answer patterns between the two aspects indicates that students did not randomly agree or disagree with the statements.

We realise that a 'perfect score' in this assessment is achieved by agreeing with all statements, which matches one of the unnatural answer patterns. However, since the answer patterns in the current data set suggest that students did not randomly agree or disagree with statements, we believe this limitation did not have a significant influence on our results. We do encourage to address this limitation in future research, for example by adding negatively worded statements to the assessment.

Our findings, that presented explicit modelling purpose (aesthetic purpose, explanatory purpose or research tool purpose) does not affect students' agreement scores for statements related to the aspects *nature of models* and *multiple models,* seem to be in contrast with previous research by Krell (2019), who found that students' expressed level of understanding does change depending on the provided modelling-purpose. However, where our results were related to the modelling aspects *nature of models* and *multiple models,* Krell (2019) focused on the modelling-aspects *testing models* and *changing models*. A possible explanation for this difference in effect is that the modelling-aspects *nature of models* and *multiple models* are inherently different from the modelling-aspects *testing models* and *changing models*. The modelling- aspects *nature of models* and *multiple models* deal with the ontological and epistemological understanding of models, while the modelling-aspects *purpose of models*, *testing models* and *changing models* focus on the use of models in science (Grünkorn et al., 2014). This means that the provided explicit modelling-purposes that reflect on the modelling-aspect *purpose of models* are more related to the aspects *testing models* and *changing models,* than to the aspects *nature of models* and *multiple models.* This inherent difference between the aspects provides a possible explanation for why an explicit modelling-purpose influences students' agreement score for the aspects *testing models* and *changing models*, but not for the aspects *nature of models* and *multiple models.*

It is noteworthy that students' agreement score for Level 3 statements was higher in this study than in previous studies that looked at students' expressed level of meta-modelling knowledge considering the aspects *nature of models* and *multiple models* (Grünkorn et al., 2014; Krell et al., 2015). In this current study students agreed with 64 percent of the Level 3 statements for the aspect *nature of models* and with 86 percent of the Level 3 statements for the aspect *multiple models*, when no explicit modelling-purpose was present. In contrast, Grünkorn et al. (2014) found that four and nine percent of the students reached a Level 3 understanding for the aspects *nature of models* and *multiple models* respectively. Krell et al. (2015) found that around 15 and 35 percent of the students preferred a Level 3 understanding for the aspects *nature of models* and *multiple models* respectively. We propose two possible explanations for this difference. The first explanation is that the concept-process models that are used in this current study possibly trigger a more advanced level of understanding than the variety of models that was used by Grünkorn et al. (2014). Because of the dynamics, such as time and movement, that are present in concept-process models, these models are considered to be more complex and abstract than for example scale models and therefore put different conceptual demands on students (Harrison & Treagust, 2000). It is possible that interpreting concept-process models requires more abstract thinking, which can trigger higher levels of meta-modelling understanding. The second possible explanation is the difference in method that was used in these studies. In this current study, students could display more than one level of understanding (i.e., students had the option to agree with more than one statement). In the study by Krell et al. (2015), students had to rank statements according to how well they thought given statements matched the given model. Students were asked to rank three statements which corresponded to the three levels of meta-modelling knowledge. This means that students had to choose only one of the statements as 'best fit'. In this current study, many students made use of the opportunity to agree with more than one statement. This means that even though it is possible that Level 3 might not be the preferred statement-level to agree with for the majority of students, this current study shows that most students are able to display Level 3 thinking when given the opportunity to agree with statements on multiple levels of understanding.

Our second research question focused on the influence of contexts showing a (sub)microscopic or a macroscopic model on students' agreement scores. Results show that students more often agreed with higher level statements (Level 2 and Level 3) when presented with (sub)microscopic models than when presented with macroscopic models. This observation is in line with the qualitative research by Jansen et al. (2019) who suggest a more advanced meta-modelling knowledge for (sub)microscopic models than for macroscopic models. The discrepancy between expressed meta-modelling knowledge for macroscopic and (sub)microscopic levels might be explained by a difference in abstractness. Macroscopic models are visible to the naked eye, which makes them less abstract than (sub)microscopic models that are not visible to the naked eye (Tsui & Treagust, 2013). The abstract thinking skills that are needed to interpret (sub)microscopic models can possibly trigger the abstract thinking skills that are needed to express the more advanced levels of meta-modelling knowledge. This possible need for abstract thinking skills resembles our previous explanation for the difference in meta-modelling knowledge between scale models and concept-process models, suggesting that this explanation of model abstractness and more advanced meta-modelling knowledge possibly explains both observations. However, when looking at the different aspects of focus, the described difference between macroscopic and (sub) microscopic models can only be found for the modelling-aspect *nature of models*, not for the modelling-aspect *multiple models*. Further research should investigate whether this is a persistent difference and whether similar differences in expressed meta-modelling knowledge between macroscopic and (sub)microscopic models can be observed for other aspects of meta-modelling knowledge.

It has to be noted that even though we provided students with a familiar biological context and took the presence or absence of various modelling purposes into account, there could still have been individual differences in the way students experienced the models and contexts that we used in the test. Khishfe (2017) argues that the familiarity with a context, the amount of exposure to a certain context, or its personal relevance, can influence students reasoning. Since our study did not take differences in exposure to a context or the personal relevance of a context into account, differences in these factors could have an influence on individual student results. Also, according to Fischer (1993) and King and Kitchener (2004), students need to be provided with contextual support in order to show an optimal level of performance. This support can be provided in different ways, such as offering students a high-level example of the skill (in this case model-based reasoning), offering students the opportunity to ask questions about the example and giving students the chance to practice a skill in a variety of settings. A study by Schwarz and White (2005) showed that familiarizing students with a context by letting them carry out real-world investigations and test hypotheses indeed promoted students' understanding of the nature and purpose of models. In this study we did not provide students with this type of support, but instead assessed students expressed meta-modelling knowledge as it was at that time. Since students came from various schools in the Netherlands, it is possible that there is a difference in the amount of contextual support that students experienced in their education. This means that our results show the average expressed meta-modelling knowledge of students in the Netherlands without receiving extra contextual support or taking personal relevance into account. Depending on the amount of support students received in class, this average expressed meta-modelling knowledge can differ from students' optimal performance after receiving contextual support. It would therefore be interesting for future research to assess students' expressed meta-modelling knowledge after receiving this type of support.

It is important to note that investigating students' current expressed meta-modelling knowledge can only be seen as a first step in improving meta-modelling knowledge. As previously described, Level 1, Level 2 and Level 3 are all valid levels of reasoning for both macroscopic and (sub)microscopic biological concept-process models. In order to stimulate model-based reasoning and scientific literacy, it is important for students to gain knowledge about, and practice with, the meta-modelling aspects and corresponding levels as described in Table 3.1. Ideally, students should be able to adjust their reasoning considering the aspects and corresponding levels to the question they have at hand. Therefore, we believe that students need to gain experience with model-based reasoning considering all aspects and levels (Level 1, Level 2 and Level 3) as described in Table 3.1.

Results from this study can be used to indicate important focus points for developing model-based learning approaches to stimulate students' expressed meta-modelling knowledge considering biological concept-process models.

Considering the aspects of focus (*nature of models* or *multiple models*), results show a significantly lower agreement score for statements considering Level 3 for the aspect *nature of models*, and considering Level 1 and Level 3 for the aspect *multiple models.* These results suggest that extra attention should be given to these levels and aspects when designing learning approaches to stimulate model-based reasoning. Also, when focusing on the different types of models ((sub)microscopic and macroscopic), our results suggest that teaching material should specifically highlight both Level 1 and Level 3 types of reasoning for (sub)microscopic models, and Level 3 type of reasoning for macroscopic models.

The test as it was used in this study has proven to offer valuable insights into students' expressed meta-modelling knowledge considering the aspects *nature of models* and *multiple models.* We believe that this test, or an extended version including all five aspects and corresponding levels, can be a useful tool in the process of assessing students' expressed meta-modelling knowledge considering biological concept-process models. It can be used to provide insight into which aspects and levels as described in Table 3.1 need extra attention when developing model-based learning approaches that focus on biological concept-process models, and to assess the effect of these teaching activities on the described aspects of model-based reasoning.

CHAPTER 4

Lesson Study as a research approach: a case-study

This chapter is based on:

Jansen, S., Knippels, M.C.P.J., & van Joolingen W.R. (2021). Lesson Study as a research approach: a case-study. *International Journal for Lesson and Learning Studies, 10*(3), 286-301. https://doi.org/10.1108/IJLLS-12-2020-0098

Abstract

Purpose: The purpose of this article is to explore the merits of Lesson Study (LS) as a research approach for research in (science) education. A lesson was developed to introduce students to model-based reasoning: a higher order thinking skill that is seen as one of the major reasoning strategies in science.

Method: Participants of the LS-team were three secondary school teachers and two educational researchers. Additionally, one participant fulfilled both roles. Both qualitative and quantitative data were used to investigate the effect of the developed lesson on students and to formulate focal points for using LS as a research approach.

Findings: The developed lesson successfully familiarized students with model-based reasoning. Three main focal points were formulated for using LS as a research approach: 1) make sure that the teachers support the research question that the researchers bring into the Lesson Study cycle, 2) take into account that the lesson is supposed to answer a research question which might cause extra stress for the teachers in a LS-team, 3) state the role of both researchers and teachers in a Lesson Study team clearly at the beginning of the LS-cycle.

Originality: This study aims to investigate whether LS can be used as a research approach by the educational research community.

4.1. Introduction

Lesson Study (LS) is known as an approach in which a team of teachers collaborates to target an area of development in students' learning by designing, teaching, observing and evaluating lessons (Fernandez & Yoshida, 2008). Studies have shown that classrooms provide powerful, practice-based contexts in which teachers learn ways to support student learning (e.g., Opfer & Pedder, 2011). Among other benefits, LS has been proven to make the teachers more aware of students' thinking processes (Verhoef & Tall, 2011) and to enhance student learning (e.g., Ming Cheung & Yee Wong, 2014).

Since LS often focuses on teacher professional development, with teachers investigating their own practice, research on LS often has focused on what teachers learn from LS (e.g., Schipper et al., 2017; Vermunt et al., 2019), or how LS can be implemented in schools (e.g., Chichibu & Kihara, 2013).

However, the cyclic nature of LS that allows for systematic refining of lessons, might not just be beneficial for addressing topics arising from the LS-team, but could also benefit the study of specific problems prominent in existing bodies of educational research. Research approaches focussing on the design of teaching and learning activities in a cyclic fashion are often labelled *Design Research* of which multiple versions exist (Bakker, 2018). Due to its cyclic nature and focus on teaching and learning, LS can be seen as a kind of Design Research. In the spectrum of Design Research, LS focuses on student learning and a strong involvement of teachers in the design process of lessons. This strong involvement of teachers allows them to integrate their experience and expertise into the design. The focus on student learning is, for instance, apparent in LS models used in the UK (Dudley, 2015) and the Netherlands (de Vries et al., 2016), in which a number of so-called *case-students* are closely observed in each lesson and interviewed afterwards. This means that, apart from student results from the whole class, detailed quotes, student behaviour and arguments are available for these case-students (de Vries et al., 2016). The large pedagogical and didactical contribution from teachers, and the amount of detailed data from the case-students resulting from a LS approach, can provide valuable insights for the research community.

A major challenge in science education is to foster students' higher order thinking skills (Miri et al., 2007). These skills are very difficult to capture and LS might especially be a beneficial approach for research focusing on this area. LS's focus on observation of student learning may help studying students' reasoning processes. Teachers' pedagogical content knowledge (PCK) can help in the design of activities that make students' reasoning abilities visible, allowing researchers to study the resulting data on student learning.

To explore whether LS has potential as an approach addressing research questions on higher order thinking, we present a case-study in which LS is used as a research approach to develop teaching and learning activities that address the higher order thinking skill of *model-based reasoning*. Model-based reasoning entails the understanding of the nature and use of scientific models as a basis for scientific knowledge. In science education research, model-based reasoning is seen as one of the major reasoning strategies that is part of scientific literacy (Windschitl et al., 2008). In this study we focus on a particular kind of model that is often used in biology: concept-process models. Concept-process models visualise biological processes such as an image showing the process of cellular respiration. These concept-process models are perceived as the most complex type of models in biology education (Harrison & Treagust, 2000). Unlike scale models or visual depictions of a certain biological phenomenon, concept-process models have a very abstract nature. They include the inherent dynamics of biological processes, such as time and movement, which are often visualized by arrows (Jansen et al., 2019). Figure 4.1 shows an example of a biological concept-process model that is used in biology education, in which the light reaction of photosynthesis is depicted. The light reaction is the first part of photosynthesis, in which energetic molecules are formed that are necessary in the process of creating glucose. The formation of glucose takes place in the second part of photosynthesis, called the Calvin cycle.

The dynamics that biological concept-process models can represent make this type of model ideal for learning about scientific processes. They can be used to explain phenomena, but also to formulate hypotheses or carry out thought experiments (Windschitl et al., 2008).

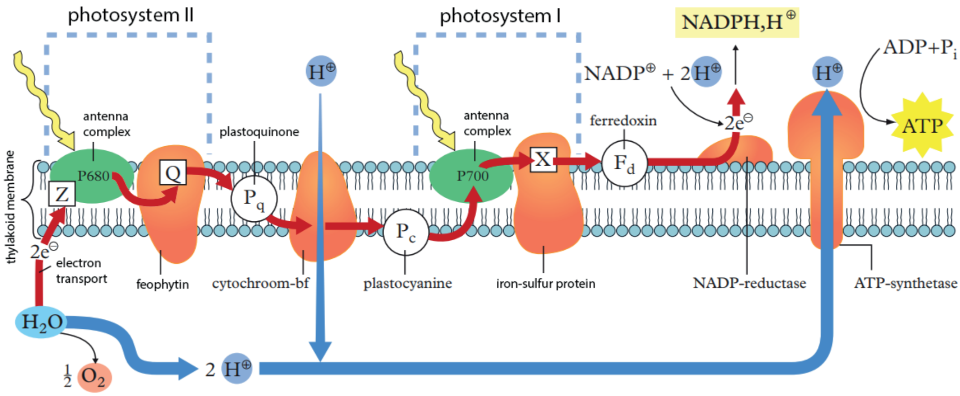


Figure 4.1. A concept-process model of the light reaction of photosynthesis. Reprinted and translated with permission from Noordhoff Uitgevers, Groningen (Brouwens et al., 2013).

4.1.1. Aspects of model-based reasoning

A framework developed by Grünkorn et al. (2014) and adapted by Jansen et al. (2019) shows five important aspects of model-based reasoning that reflect on understanding and using models in science (Table 4.1). The five aspects within this framework are: *nature of models, purpose of models, multiple models, testing models* and *changing models.*

The aspects *nature of models* and *multiple models* include the way models are used to describe and simplify phenomena. *Nature of models* focuses on the extent to which the model can be compared to the original, whereas *multiple models* refers to the fact that various models can be used to represent the same original. Both aspects show that models are often simplified and emphasize only those elements that are important to explain a certain key idea (Harrison & Treagust, 2000). The aspects *purpose of models, testing models* and *changing models* focus on the use of models in scientific practices. These include testing hypotheses, making predictions about future events and communicating ideas (Grosslight et al., 1991). With the aspect *purpose of models* the framework focuses on aims that can be met using a certain model. *Testing and changing models* describe the way a model is being validated and stress the fact that models are by definition temporary and changeable. For all these aspects, up to four levels of understanding have been determined, ranging from an initial level of understanding to an expert 'scientific' level of understanding (Level 3).

Table 4.1. Framework to assess students' understanding of biological models. The left column shows the five aspects that are important when reasoning with biological models. For each of these aspects up to four levels of understanding have been defined, ranging from an initial level of understanding to an expert level of understanding (Jansen *et al., 2019*).

	Initial level	Level 1	Level 2	Level 3
Nature of models		- Model as copy - Model with great similarity - Model represents a (non-) subjective conception of the original **- Displays a process, its components and how they are related**	- Parts of the model are a copy - Model as a possible variant - Model as focused representation	- Model as hypothetical representation
Purpose of models		- Model for showing the facts **- Model for showing events**	- Model to identify relationships	- Model to examine abstract/ concrete ideas
Multiple models	- All models are the same - Various models of different originals - Only one final and correct model	- Different model object properties	- Focus on different aspects	- Different assumptions
Testing models	- No testing of models **- Perceiving schoolbooks or their authors as authorities providing absolute truth**	- Testing of material - Testing of basic requirements	- Comparison between original and model - Comparison and matching of original and model	- Testing hypotheses - Testing of hypotheses with research designs
Changing models	- No reason for alterations - Alteration of how different originals are represented	- Alterations to improve the model object - Alterations when there are errors in the model object - Alterations when basic requirements are not met	- Alterations when model does not match the original - Alterations due to new findings about the original	- Alterations due to findings from model experiments **- Alterations when the focus of the model shifts to a different aspect of the process**

4.1.2. Aim of the study

The aim of this study is to explore whether LS can be a suitable research approach for answering questions on higher order thinking skills. This case-study demonstrates the application of LS as a research approach to develop so-called *key activities* that explicitly introduce students to the aspects and levels of model-based reasoning as described in Table 4.1. In this case-study we will focus on whether the use of LS answers the research question: how can students successfully be familiarized with important aspects of model-based reasoning? As a second step we will evaluate what we learned from this case-study on using LS as a research approach, and formulate recommendations for using LS in answering research questions on higher order thinking skills. This leads to the following two research questions:

> *4.1: To what extend does the developed lesson successfully familiarize students with important levels and aspects that are associated with model-based reasoning?*
>
> *4.2: How do teachers and researchers experience using LS as a research approach?*

4.2. Method

4.2.1. Participants

Three biology teachers (18, 30 and nine years of experience, one male, two female) from the same secondary school, two researchers (second and third author) and the first author who is both a researcher and a secondary school biology teacher with eight years of experience participated in a LS team. The lesson was performed and observed in two 11th grade pre-university biology classes from the Netherlands. In total 34 students (16–18 years old, 18 female, 16 male) engaged in all scheduled activities.

4.2.2. RQ4.1: Pre- and Post-Test

Online pre- and post-tests were developed to determine up to which level the lesson familiarized students with the three aspects of model-based reasoning (*nature of models, purpose of models* and *multiple models*). Both tests contained the same nine open-ended questions, where students had to formulate a definition of a model in biology, and answer questions relating to the aspects *nature of models* (two questions), *purpose of models* (three questions) and *multiple models* (three questions). An example of a question relating to the aspect *purpose of models* is: “Before a biological model is made, the creator of the model thinks about what the model will be used for. Indicate a possible purpose of the model below”. A translated list of all questions is available in the supplementary material. The pre-test took place in the biology lesson preceding the developed lesson and the post-test in the biology lesson following the developed lesson.

4.2.3. RQ4.1: Interviews – Case-Students

The six case-students, three for each version of the lesson, were interviewed after the lesson using a semi-structured interview scheme. The questions as proposed by de Vries et al. (2016) were used: students were asked about what they liked about the lesson; what they learned from the lesson; what they thought worked well in the lesson; and what they would change about the lesson if this lesson would be taught again to a different class. Interviews were recorded and lasted 5-10 minutes.

4.2.4. RQ4.2: Interviews - Teachers

The three teachers who participated in the LS-team were interviewed after the completion of the LS-cycle using a semi-structured interview scheme to evaluate what the main focal points are when using LS as a research approach. Interviews were recorded and lasted approximately 40 minutes. The interview questions related to the expectations the teachers had before starting the LS-cycle and to what extent these expectations were met; what they thought went well during the LS-cycle and what not; what they learned from participating in a LS-cycle; the extent to which they applied what they learned to other lessons or their teaching; and whether they expected to keep on using what they had learned in the long term.

4.2.5. RQ4.2: LS-meetings

The LS-cycle started with an introduction on model-based reasoning by the researchers to the teachers in the LS-team. A 45-minute lesson was then designed in three two-hour meetings within a timeframe of two weeks. The LS-team evaluated both the designed and adapted lesson in a one-hour meeting. All meetings were audio-recorded.

4.2.6. Data Analysis

Student answers on the pre- and post-test were coded using the three aspects of interest and their corresponding levels as described in the framework from Grünkorn *et al*. (2014) as codes. Possible student answers for the aspect *purpose of models* are as follows.

> *Question: Before a biological model is made, the creator of the model thinks about what the model will be used for. Indicate a possible purpose of the model below.*
>
> *Level 1: To show the different parts of a plant*
>
> *Level 2: To indicate what relationships are present between this process and other processes*
>
> *Level 3: To display the process of fertilization, after which the researcher can use the model to do research on the process*

Fifty percent of the answers were coded by a second independent coder, resulting in a Cohens' Kappa of 0.69 for *nature of models;* 0.87 for *purpose of models* and 0.63 for *multiple models*. Student answers in the *audio recordings* were tagged when utterances related to aspects that were learned from the lesson. Tagged answers were grouped according to the three aspects of focus and three levels of reasoning. Student *material* was tagged for utterances relating to the three levels for each of the three model-based reasoning aspects.

To learn from the experience of using LS as a research approach, the audio recordings from both the teacher interviews and the LS-meetings were tagged for utterances relating to elements that worked well and for elements that needed improvement. Audio tags were grouped into these two categories.

4.3. Results

4.3.1. Lesson Design – The Design Process

After the theoretical introduction by the researchers, the first LS-meeting was used to decide on the curriculum topic for the lesson and the models to be used in that lesson. The second LS-meeting focused on formulating key activities that let students reflect on the aspects within the framework (Table 4.1) that the team wanted to get students acquainted with. During the third LS-meeting the LS-team decided on the three case-students that would be observed in detail during each performance of the lesson and on predicting the learning behaviour of these students.

In selecting case-students the LS-team made use of the expertise of the teachers and their knowledge about the students. Since the levels in Table 4.1 represent an increasing degree of difficulty, the LS-team assumed that students who were able to reason on Level 3, would also be able to reason on Level 1 and Level 2. Therefore, teachers were asked to define for every student if they thought the student would have a high chance, an intermediate chance or a low chance of reaching Level 3 for the aspects as described in the framework. They also indicated which of these students would be explicit in their arguments, making it easier to follow their way of reasoning during the lesson. Students were placed in homogenous groups of four students, based on this classification. From three groups a case-student was selected. For each case-student an observation scheme was created, listing their predicted behaviour during each phase of the lesson. For each case-student a back-up student was chosen and an observation scheme was formulated, in case one of the selected students would not attend the lesson. Case-students were observed by members of the LS-team, using the observation scheme.

The reason students were placed in homogenous groups was mostly pragmatic. Each observer was stationed next to a group of students of whom the teachers expected a certain behaviour. The observer could remain seated next to this group of students and observe the back-up case-student in case the selected case-students did not attend the lesson.

One of the teachers from the LS-team taught the lesson. Discussions that took place in the student groups containing case-students were audio recorded and all student work was collected. After teaching the lesson for the first time to one of the biology classes, the lesson was discussed with the LS-team and improvements were formulated. The adjusted lesson was taught one week later by the same teacher in a different biology class.

4.3.2. Lesson Design – Aspects of Focus

The LS-team decided to focus on three of the five aspects listed in Table 4.1, and to design a key activity for each of these three aspects. The teachers indicated that time was an important factor to take into account. The lesson duration of 45 minutes was considered to be too short to properly introduce all five aspects. Since the aim of this study was to introduce aspects that are important when reasoning with *existing* biological concept-process models, the researchers in the LS-team explained that the aspects *nature of models, purpose of models* and *multiple models* would be the aspects of choice when creating the lesson. These aspects are central to understanding given models and are important when reasoning with these models, such as the ones students encounter in their textbook. The aspects *testing models* and *changing models* are of importance when a model is *created, tested or modified*.

4.3.3. Lesson Design - Pedagogical Choices

The LS-team made various pedagogical choices considering the design of the lesson. These choices were mostly based on teachers' pedagogical knowledge and experience and discussed with the researchers on the team, who searched for literature backup.

The teachers decided on photosynthesis as the subject of the lesson, since many models about photosynthesis are available for educational contexts. Also, this topic was recently taught in class, and, according to the teachers, this allowed for focussing on the model-based aspects and not on the content domain. This choice is in line with literature on this topic, showing that students need domain knowledge before they are able to create their own mental model of a process (e.g., Cook, 2006) or interpret given scientific models (e.g., Tasker & Dalton, 2007).

In order to engage students with the lesson and theory about the aspects of model-based reasoning, the teachers in the LS-team wanted students to work with these aspects themselves before explaining the theory. According to the teachers, just explaining or showing the theory to the students would put the students in 'consumer-mode'. An inductive approach, where students have to think about the theory themselves first, would engage the students and make them curious for answers. This choice is backed up by research showing that inquiry-based learning stimulates scientific reasoning and helps students to gain confidence in their scientific abilities (Gormally et al., 2009).

To provide insight in student thinking, the teachers decided that the developed key activities should stimulate students to work together and talk out loud during the lesson. Research shows that talking out loud is not only beneficial for providing insight in student thinking, but also promotes student thinking about what they understand and what not, thereby improving metacognition (Tanner, 2009). Also, working in groups can improve student performance in general and aid in learning (Smith et al., 2009).

4.3.4. Lesson Design – Resulting Key Activities

The pedagogical and didactic choices that were formulated by the LS-team were incorporated into three key activities.

In **key activity 1** students focused on the aspect *multiple models*. Students were asked to individually name differences between four models that showed the same biological process (photosynthesis) (Figure 4.2). These differences were then shared in groups of four students, after which the group categorized the differences. The teacher then linked these categories to the levels as represented in the framework (Table 4.1).

In **key activity 2** students matched aims of a model to the four models of photosynthesis. This activity corresponds to the aspect *purpose of models.* The aims were provided by the teacher and were formulated according to the three levels as described by the framework (Table 4.1). In order to stimulate discussion and have students substantiate their choices, they were only allowed to match an aim with one of the models when everyone in their group agreed on this choice. The teacher then discussed the results and explained how the aims related to the three levels as described by the framework.

In **key activity 3** relating to the aspect *nature of models,* students were assigned to one of the four models. Students had to formulate the choices that the creator of the model had made to meet the aims. They also indicated which components of the model were drawn in a true to nature way, and which were not. Afterwards the teacher linked students' choices to the levels as described by the framework, and explained how these choices relate to these levels. Figure 4.3 summarizes the design process and shows the contributions of both teachers and researchers to the final lesson design.

After teaching the lesson for the first time to one of the biology classes, the lesson was discussed in the LS-team. Only a minor adjustment was made, the four models of photosynthesis were numbered (1-4) before teaching the adjusted lesson.

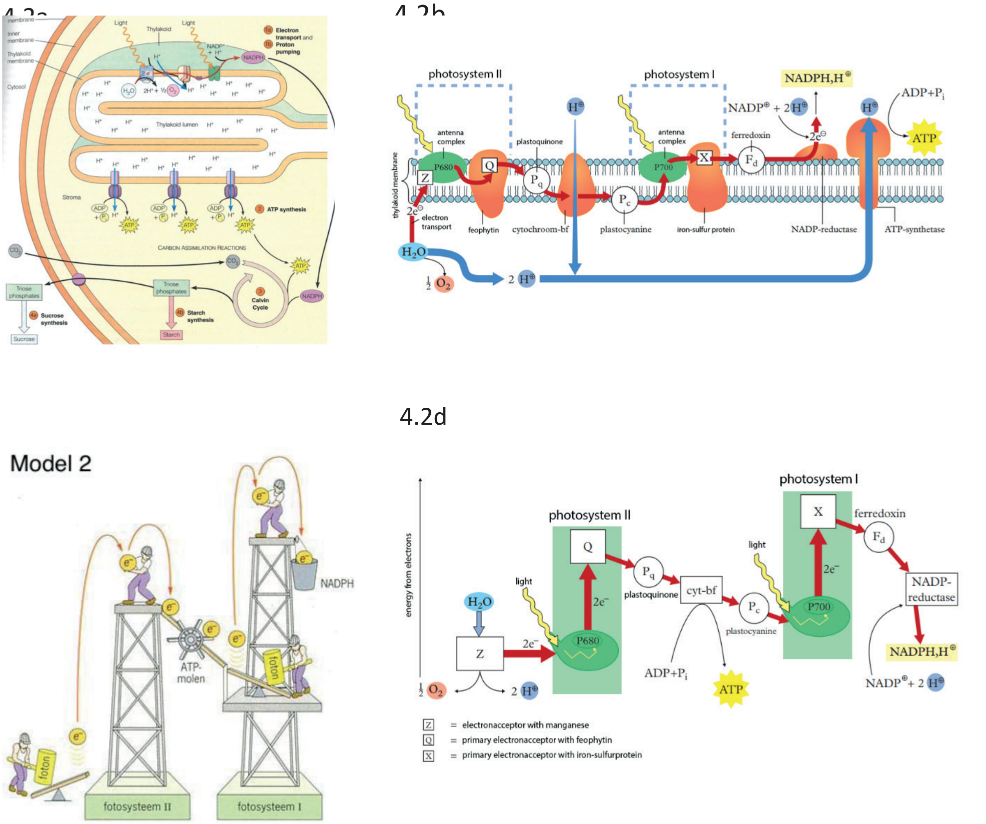


Figure 4.2. The four models of the light reaction of photosynthesis that were used in the developed lesson. All models show the same process, but have a different appearance due to differences in emphasis or choices of the creator. Figure 4.2a focuses on the light reaction (first part of photosynthesis), but also shows the connection to the Calvin Cycle (second part of photosynthesis). Figure 4.2b zooms in on a part of the thylakoid membrane with the electron transport chain, leaving out the connection to the Calvin cycle. Figure 4.2c focuses on the energetic state of the electrons and the role of photons in this process. Figure 4.2d shows resemblance with Figure 4.2b, but places more emphasis on the energetic state of the electrons and the proteins that are involved, leaving out the thylakoid membrane. All figures are reprinted with permission from Pearson Education, San Francisco (Figure 4.2a and 4.2c) and Malmberg, 's Hertogenbosch (Figure 4.2b and 4.2d (both translated with permission from Dutch to English)) (Brouwens et al., 2013; Campbell & Reece, 2002).

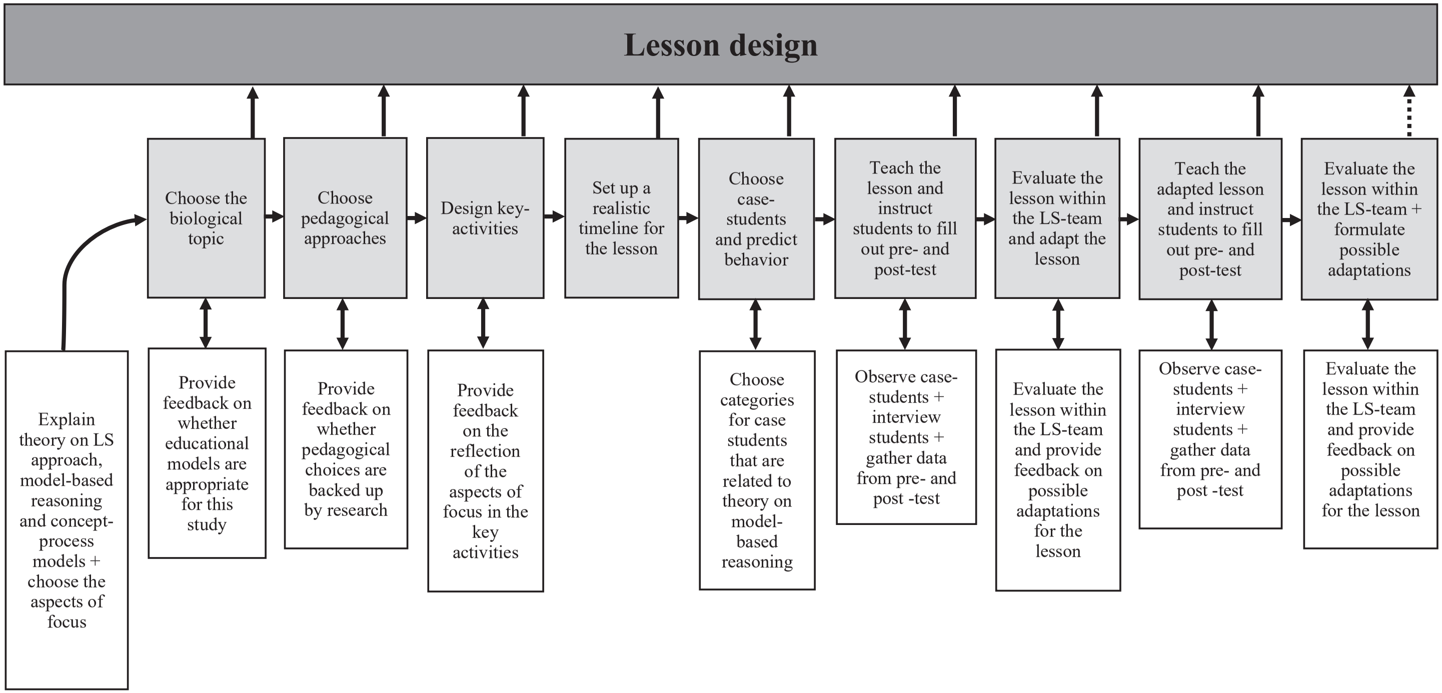


Figure 4.3. Contributions of the teachers and the researchers to the lesson design. The contributions from the teachers are visualized in light grey. The contributions from the researchers are visualized in white. The arrows show interactions between the researchers, teachers and the lesson design. The dotted arrow shows a possible adaptation moment to the lesson design. In this case-study these possible adaptations have been discussed, but have not been applied to the lesson design since the lesson would not be taught a third time. The figure can be read as a timeline from left to right.

4.3.5. Influence on students' reasoning

Even though the subject of this lesson, model-based reasoning, is not part of the curriculum, all six case-students mentioned during the interviews that they enjoyed the lesson and that they would like to learn more about this subject. The following quote shows how one of the case-students felt about this lesson (translated from Dutch to English).

> *LS1B1: I thought it was very interesting. It was a different way of looking at the theory. When you learn to look at the theory in this way, you will understand it better. I really feel that way.*

Based on the interview data and lesson recordings, we obtained insight in the learning process of the six case-students. Considering the aspect *multiple models*, all case-students were, individually during the interview or together within their group, able to name various kinds of differences between models of the same process. For all case-students, the formulated differences related to multiple levels of reasoning within the framework (Table 4.1). For instance, when asked about the kinds of differences found, a student answered:

> *LS1A1: Well, we formed a group [of differences] about content, so what is visible in the image, or how much is being shown. In one of the pictures for example you can also see the Calvin cycle and in the other picture you cannot. And we have [a group of differences] about what the image is meant for. For example, the one with the small guys in it [Figure 4.2c]. In that one the focus is only on the levels of energy of the electrons. And then we also have [a group] with visual differences, which is about the fact that some images have been drawn in a more realistic way than others.*

In this case the group of differences about 'content' and the group of differences about 'what the image is meant for' both relate to Level 2 for the aspect *multiple models*, since they address the differences in focus between several models. The group with 'visual differences' relates to Level 1 for the aspect *multiple models*, since it addresses different model object properties.

Considering the aspect *purpose of models,* all case-students were able to match different aims with different models of the same process and explain why they matched a certain aim. The various aims related to the different levels within the framework for the aspect *purpose of models*. The following student discussion shows how one of the aims was matched with a specific model:

> *L1B1: Shall I read the aim out loud?*
>
> *L1B2: Yes, that way we can think about it together*
>
> *L1B1: To show that electrons are released when water is splitted.*
>
> *L1B3: I think that's this one, because it clearly shows that water is splitted [pointing at the model in Figure 4.2d].*
>
> *L1B4: Yes, but you can see that in this model too. And in this one [pointing at the models in Figure 4.1a and Figure 4.2b]!*
>
> *L1B2: Yes, but I think it should be the one where the focus of the model is on splitting water.*
>
> *L1B1: Well, this model really emphasizes the presence of electrons, you can literally see two electrons appearing [pointing at the model in Figure 4.2d].*
>
> *L1B4: Yes, but you can also see that in the other models*
>
> *L1B3: Yes, but the emphasis is less on the process of splitting water.*
>
> *L1B1: Ok, so let's go with this one, because the emphasis of the model is on the splitting water part and on what the electrons do, the other parts of the process are less prominent [points at the model in Figure 4.2d].*

Considering the aspect *nature of models*, which was the subject of key activity 3, all case-students showed that they were able to explain that the creator of the model makes choices in order to meet the prospected aim of the model. When asked about these choices, all case-students referred to aspects being left out or being put in to focus on a certain part of the process. This refers to Level 2 of the aspect *nature of models*. The student quote below shows an example of a student quote where the choices of the creator of the model are related to the aims that could be met using this model:

> *L1C1: If you take it literally, I don't think that there is someone using the hammer in real life.*
>
> *L1C4: It is very schematic*
>
> *L1C1: I think it's a choice to meet the aim by only showing a part of the reaction*
>
> *L1C3: Yes, simplifying it*
>
> *L1C4: Yes, focusing on a specific part of the reaction, showing that part.*

The aim of the model in this case was “to show that energy is necessary to let the light reaction take place”. The students explain that by simplifying the model, the focus is on that part of the process.

Considering the pre- and post- test, no significant differences in students’ level of reasoning for each of the three aspects of focus were found. Table 4.2 summarizes the changes in student levels on the different aspects of model-based reasoning.

4.3.6. Teachers’ experience

In the teacher interviews, all three teachers mentioned that the theory about reasoning with biological models was considered to be an eye-opener. They mentioned that this knowledge did not only affect the design of the lesson, but also led to a different way they currently teach about models in other lessons and want to keep on teaching about models in the future. The quote below shows how the introduction to this theory changed the teacher’s view on model-based teaching, causing him to intend to implement this theory in his current teaching.

> *T2: Making the role of models more explicit, that is something I will handle differently from now on. I would assume it to be less clear for students. And I think I would start with that when we use models in lower secondary education, saying “this has been visualized in this way, which is a choice of the creator of the model”.*

Table 4.2. Comparison of the pre- and post-test. The three aspects of interest are shown in the left column. The table shows for each of the two biology classes how many students (n) decreased in level of reasoning, showed no change in level of reasoning, or increased in level of reasoning.

	Biology class 1 (n=16)			**Biology class 2 (n=18)**		
	Decreased in level (n)	**No change in level (n)**	**Increased in level (n)**	**Decreased in level (n)**	**No change in level (n)**	**Increased in level (n)**
Nature of models	1	13	2	3	11	4
Purpose of models	2	13	1	4	13	1
Multiple models	1	14	1	6	10	2

A

A s shown by the following quote, the participating teachers considered the relation between theoretical aspects of models and their practical application in the lesson during the meetings in which the lesson was developed extensively.

> *T2: In most biology lessons we use models as an illustration, to explain a certain biological phenomenon. In this lesson the model itself will be the subject of the lesson. I think we need to let students think about the nature of a model and the differences between multiple models of the same biological process.*

All three teachers were positive about being part of the LS-team. Apart from the fact that the theory about model-based reasoning was experienced as a welcome new insight by all teachers, they felt that the experience brought the team of teachers closer together, and they were proud of what they had achieved during the LS-cycle.

> *T3: It's a good thing to critically discuss how to teach students about a certain subject. Together you will hear and see more perspectives than when you develop a lesson by yourself. We formulated a goal for the lesson and discussed how we could achieve the desired results. And everyone [in the LS-team] has different ideas about that. These are probably all good ideas, but because you discuss them together, the final idea will be different from your own initial idea. And because you critically look at the ideas together, the final idea will be better.*

All teachers also indicated that they would consider using LS again when developing lessons. However, they also mentioned that the LS-cycle took more time than expected and that they would therefore not see themselves participating in multiple LS-cycles in a short period of time.

> *T1: We could definitely use this method [LS] again, but perhaps I would prefer developing a lesson series instead of a single lesson. It really costs a lot of time. I look at LS as a good method to develop complete projects for example.*

Teachers were aware of the fact that the lesson was supposed to answer a research question and therefore had to lead to results that could be measured and that could be compared with the expected student behaviour that the LS-team formulated:

T2: I liked thinking upfront about what actions a certain student would undertake. It really makes you think about that specific student and whether you can predict for this student what will happen. That influences the way you teach as well. You start to behave in a certain way, because you really want things to work out the way you thought they would. And you especially want to make sure that the results could be measured.

The teacher who taught the lesson mentioned one downside about using LS as a research approach specifically. In his opinion, the script and the associated timeframe were problematic factors during the lesson.

T2: Teaching the lesson was such a strict process for me! We agreed on a certain amount of time per element within the lesson. That is really different from the way I usually teach, where I am more concerned with how the students respond, and where I adapt my teaching to their response. Now I had to do exactly what it said in the script, which meant I kept on looking at my watch. I really struggled with that, because I was afraid that the students wouldn't get the point if I wasn't able to finish all elements within that lesson. During a normal lesson I would think, that's ok, and I would continue with the theory the next lesson. Now it's just one single lesson and there are observers and we do a test, so everything needs to be finished. That caused a lot of pressure, it felt unnatural.

4.4. Discussion

While LS originally mainly focuses on teacher professional development, we used LS as a form of Design Research (Bakker, 2018) to develop a lesson that addresses a problem that arises from the existing body of research and relates to higher order thinking skills. In this case-study we followed the LS-cycle as described by de Vries et al. (2016) to design a lesson containing three key activities that introduce main aspects of model-based reasoning (Table 4.1). We combined pedagogical and didactic knowledge and experience of teachers with theoretical knowledge to develop a lesson that answers the researchers' question on how to address model-based reasoning as a higher order thinking skill in class.

The influence the teachers and researchers had on the design of the lesson in this case-study differs from the influence teachers have in a regular LS-cycle (Figure 4.3). In this study, the researchers were the ones introducing the subject for the lesson (model-based reasoning). As in a regular LS-cycle, the teachers then developed the lesson and used their knowledge and experience to make pedagogical and didactic choices. However, different from a regular LS-cycle, the researchers reflected on whether the teachers' choices were in line with theory from literature and whether the developed activities reflected the subject that the researchers had intended. The researchers were also responsible for developing a pre- and post-test to determine whether the lesson affected students' level of model-based reasoning.

Considering our first research question, we found that after the lesson all case-students were capable of reasoning on multiple levels for the aspects *nature of models, purpose of models* and *multiple models*. These results indicate that all case-students understood the meaning of the three aspects of model-based reasoning, and were able to work with these aspects on different levels of reasoning.

The pre- and post-test showed that student levels of reasoning did not significantly change for any of the three aspects of focus. However, the open structure of the questions in the pre- and post-test invited students to answer on their preferred level of reasoning. This means that even when students were capable of reasoning on multiple levels as shown in Table 4.1, the test offered the possibility to only answer on the level they preferred. Therefore, the pre- and post-test probably indicate the students' preferred level of reasoning instead of their highest capable level of reasoning. Despite this lack of increase in students' preferred level of reasoning, the qualitative data showed that all case-students were able to reason on multiple levels for each of the three aspects of focus. Considering our first RQ, we therefore conclude that the results indicate that the developed lesson successfully familiarized students with main aspects of model-based reasoning. However, future research should focus on developing lessons to deepen students' understanding on this subject, and on developing a test to assess the students' capability of reasoning on all levels separately for each of the main aspects of model-based reasoning.

Considering RQ2, using LS as a research approach was appreciated by the teachers. The teachers enjoyed being part of the LS-team and thought it was a productive way to develop lessons, stimulate creativity and increase team-spirit. All teachers mentioned that the theory about model-based reasoning was an eye-opener to them, which not only influenced their own way of reasoning with models, but also the way they intended to work with models in their future lessons.

However, results from the teacher interviews show that teachers experience one downside of being part of the LS-team, time. In this case the factor time did not only apply to how long it took to develop the lesson, but also to the strict schedule that was set up for the lesson. The teacher who taught the lesson reported pressure on performing the lesson precisely according to this schedule, as he felt this was necessary to answer the research question of the researchers.

Since we as authors fulfilled the role of researchers, it was not possible to objectively investigate the experience of the researchers in this case-study. However, we can say that as researchers we felt positive about being part of the LS-team and about using LS as a research approach. Since the teachers designed the lesson, making pedagogical and didactic choices, the role of the researchers was mainly to inform the teachers about the theoretical background and check whether the choices that the teachers made were backed up by research. We found that this approach, in combination with the teachers' important role in observing and evaluating the lesson, led to increased ownership for the teachers. Also, as researchers we felt that the practical and pedagogical knowledge and experience from the teachers added value to the developed lesson, while the theoretical knowledge that we shared with the teachers added value to the teachers' way of teaching. In our experience this exchange in knowledge improved the lesson design and served as an example of a possible way to sustainably incorporate theoretical knowledge from the educational research community into the classroom.

As mentioned in the method section, the LS-team consisted of three teachers, two researchers and a third researcher who was also a teacher. This third researcher fulfilled tasks both as a researcher and as a teacher, functioning as a bridge between the researchers and the teachers, contributing both theoretically and practically. Future research is necessary to find out whether the separation in tasks as described in Figure 4.3 also works well when the LS-team does not contain a member who is both a researcher and a teacher.

Our results suggest a number of focal points that should be taken into account when using LS as a research approach. First of all, it is important to make sure that the teachers support the research question that the researchers bring into the LS-cycle and that they are invested in designing lessons that answer this research question. This differs from the regular LS approach, where the teachers are the ones who decide on the subject of the lesson, making them naturally more aware of the need to work on this subject. To increase the teachers' support in answering the research question, we would therefore advise to extensively discuss the subject that the researchers bring to the LS-cycle. Also, as shown in previous research (e.g., Wolthuis et al., 2020), exploring possibilities to facilitate teachers and making sure that they have time to work on designing the lessons can help to increase teachers' investment.

Second, it is important to take into account that the fact that the lesson is supposed to answer a research question can cause extra stress for the teachers. As shown in this case-study, teachers could feel like they have to perform well, because they would otherwise hinder the research, or that not performing well would place an extra burden on the researchers who observe the lesson. Adding extra cycles to the LS approach might solve this problem. That way both the lesson and the way of teaching can be reviewed multiple times, making the teachers more comfortable with teaching the lesson. In this case-study, the teacher who taught the lesson indicated that he already felt more comfortable the second time he taught the lesson.

Third, it is important to be clear about the role of both the teachers and the researchers in the LS-team. That way both the teachers and researchers share responsibility for the lesson plan. As shown in this case-study, the teachers' sense of ownership considering the lesson design led to a product that was created by the whole team, of which they were proud. This is in line with results from Dudley et al. (2019), who show that teachers in a LS-team experience a high degree of ownership while collectively trying to understand how students navigate curricular pathways and pedagogies.

This case-study provides an exemplar for how LS can be used as a research approach. We believe LS is a promising approach to bring the pedagogical and didactic knowledge and experience from teachers, and the theoretical knowledge from the educational research community together and might thereby contribute to bridging the gap between theory-driven research and educational practice.

Supplementary Material

List of translated questions from the pre- and post-test

1. Models are often used in biology. Below you find three examples of biological models. Can you formulate a definition for a biological model? [three models: a scale model of a human eye, a model (drawing) of a cell, a model (drawing) of the process of pollination]
2. Every biological model is made with a certain purpose. Name two or three reasons (purposes) for creating a model of a biological phenomenon.
3. Before the model below was made, the creator of the model first decided on the purpose that this model would serve. Indicate for the model below what you think is the purpose for which this model was created. [model of the process of pollination]
4. To what extent does this model correspond to the original, real world situation? Explain your answer.
5. To meet the purpose as described in question 3, the creator made specific choices while creating this model. Describe a minimum of three choices that were made by the creator of the model to meet this purpose.
6. Can this model also be used for a different purpose? If so, give one or two examples of such purposes.
7. Often multiple models about the same biological process exist. What could be a reason for the fact that multiple models about the same process exist?
8. When multiple models about the same biological process exist, is in that case per definition one model better than the other? Explain your answer.
9. Below you find two models about the same biological process. Choose between the following statements and explain your answer
   a. The existence of both of these models is important
   b. One of the models is better/more useful than the other
   c. It would be good to combine both models and create one ultimate model

   [two models about the process of protein synthesis, both with a different focus: one model focusing on the binding of the anticodon on tRNA to the codon on mRNA, and one model focusing on the movement of ribosomes along the mRNA]

Questions relating to the aspect:

Nature of models: 4, 5

Purpose of models: 2, 3, 6

Multiple models: 7, 8, 9

CHAPTER 5

Fostering students' meta-modelling knowledge about biological concept-process models

Jansen, S., Knippels, M.C.P.J., & van Joolingen, W.R.

Abstract

The creation and application of models is one of the core practices in science, making it important to foster students' knowledge on models and the use of models in science (i.e., meta-modelling knowledge). A type of model that is often used to illustrate biological phenomena is a concept-process model. This dynamic type of model has great potential for learning about scientific practices and the nature of science, but is most often only used by teachers to illustrate biological phenomena. Therefore, students often find it difficult to reason with this type of models. This study employs Lesson Study (LS) for designing teaching and learning activities for secondary biology education, focussing on model-based reasoning with biological concept-process models in order to foster students' meta-modelling knowledge. The LS-team, consisting of three teachers and a researcher, made pedagogical and didactical choices when integrating suggested activities from literature in the design of two lessons (90 minutes each) to 1) stimulate students' visual literacy, and 2) support model-based reasoning via drawing. Both lessons consisted of multiple key activities, e.g., giving meaning to colours and arrows and creating a model of a biological process. Pre- and post-tests and data from student interviews showed that the lessons contributed to improving students' (n=61, pre-university level, 16-18 years old) expressed meta-modelling knowledge for several aspects relevant for model-based reasoning, such as the nature of models and the reason multiple models of the same phenomenon exist. In this study, LS bridged theory and practice and showed how teachers not only used the presented theory for developing these two lessons, but implemented the theory sustainably into their teaching.

5.1. Introduction

An important goal in science education is that students acquire insight in scientific practices and the role of science in society. Such insights are commonly seen as belonging to *scientific literacy* (DeBoer, 2000; Hodson, 2014). Since the results of scientific research, for example on the subject of vaccines and herd immunity, has penetrated our society to such an extent that functioning without basic scientific knowledge is very hard, it is important that the various practices of science are part of the science curriculum for all students.

One of the core practices in science is the creation, testing, and application of scientific models (Gobert & Buckley, 2000). Scientists can create models to describe or simplify phenomena, present scientific findings, facilitate the testing and revision of scientific theories (Svoboda & Passmore, 2013) or visualize phenomena that are not visible with the naked eye (Francoeur, 1997). When learning about science, students have to learn about models, as well as their creation and use, in order to learn and understand models and the science behind them (Grosslight et al., 1991). Modelling has therefore been defined as one of the core practices that can improve students' understanding of the nature of science (García-Carmona & Acevedo-Díaz, 2018; National Research Council, 2012).

Since students' understanding of models and modelling is strongly related with their experience with creating and using them, it is important for teachers to engage students in modelling practices on a regular basis (e.g., Krell et al., 2012; Treagust et al., 2002). When students are creating models, they actively engage in scientific practices (Bierema, Schwarz, & Stoltzfus, 2017), supporting subject matter expertise and experience the practices of building and evaluating scientific knowledge (Lehrer & Schauble, 2006; Schwarz & White, 2005). Also, by defining science as a process of creating, testing and applying models, students understand that scientific knowledge is a human construct in which models are used to explain and predict real world phenomena (Gilbert, 1991). This means that having a basic understanding of the scientific nature of models can help students in developing and evaluating explanations of the real world.

In the context of biology education, concept-process models are often used to describe processes and phenomena. Even though the dynamic nature of these models makes them suitable for teaching students about scientific practices such as formulating hypotheses, teachers often use models only to describe biological phenomena, without discussing possible scientific practices (Windschitl et al., 2008). Students therefore often have trouble understanding the scientific nature of concept-process models (Harrison & Treagust, 2000). This study aims to foster students' knowledge on models and the use of models (i.e., meta-modelling knowledge) for biological concept-process models specifically. Research has shown that both drawing models and stimulating visual literacy aids in fostering students' model-based reasoning and meta-modelling knowledge (e.g., Ainsworth, Prain, & Tytler, 2011; Gijlers, Weinberger, van Dijk, Bollen, & van Joolingen, 2013; Quillin & Thomas, 2015). This study will therefore integrate these two aspects when designing teaching and learning activities to foster students' model-based reasoning and meta-modelling knowledge considering biological concept-process models.

5.2. Theoretical Framework

5.2.1. Concept-process Models

Models come in many kinds and representations. In biology education, models range from scale models (e.g., a torso showing the placement of organs) to abstract models representing biological processes (e.g., a representation of photosynthesis or cell division). The latter type of models is often labelled as *concept-process models* (Harrison & Treagust, 2000). In contrast to scale models or visual depictions of a certain biological aspect, concept-process models show the inherent dynamics of biological processes, such as time and movement (Jansen et al., 2019). Because of their abstract nature, concept-process models are perceived as the most complex type of models in biology (Harrison & Treagust, 2000). The inherent dynamics of concept-process models makes them very suitable for learning about scientific practices and the nature of science. They can be used to explain how a certain phenomenon works, and to formulate hypotheses and reason about them using thought experiments. Figure 5.1 shows an example of a concept-process model in which the 'light reaction of photosynthesis' is depicted. It shows how the energy from light is used by plants to generate Adenosine Triphosphate (ATP), a molecule that stores energy and is used by the plant in the formation of glucose molecules.

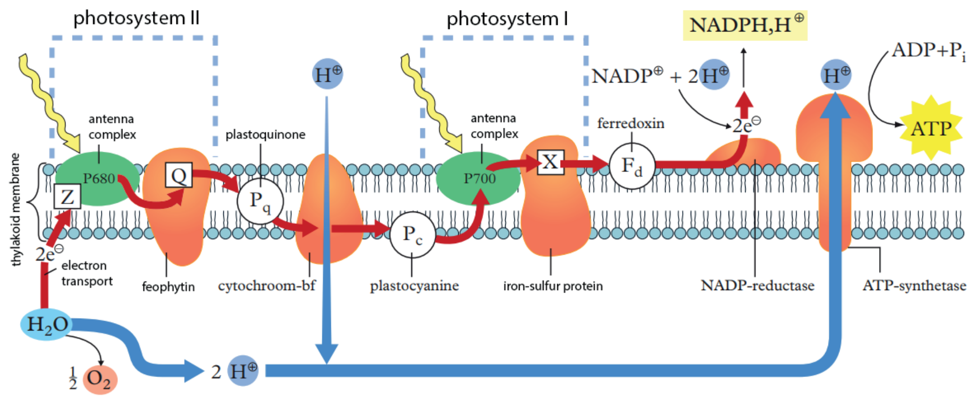


Figure 5.1. A concept-process model of the light reaction of photosynthesis. The model shows the thylakoid membrane of a chloroplast in a plant cell. Reprinted and translated with permission from Noordhoff Uitgevers, Groningen (Brouwens et al., 2013).

5.2.2. Model-based reasoning and meta-modelling knowledge

Two concepts are often used when talking about working with models in education: model-based reasoning and meta-modelling knowledge. Meta-modelling knowledge can be defined as the knowledge about the nature and purpose of models in general (Schwarz & White, 2005). This means that students are aware of the scientific practices that are necessary to generate, test and revise models (the nature of models), and that models can serve various purposes (e.g., describing or simplifying phenomena, testing hypotheses). Considering model-based reasoning, there seems to be no consistent definition available in literature. For example, Treagust et al. (2002) talk about model-based reasoning as "the application of models beyond a descriptive nature", where models are used in a quantitative or interpretive fashion. Kouw (2015) describes model-based reasoning as "the various ways in which models and modelling are intertwined with scientific practice and technological development", and Quillin and Thomas (2015) refer to model-based reasoning as "a type of problem solving that enables analysis of complex and/or abstract concepts". For the purpose of this paper, we refer to model-based reasoning as the practice of solving problems using models. Since the ability to adequately reason with models shows an understanding of both the potential and limitations of models, we believe that model-based reasoning and meta-modelling knowledge are inherently intertwined.

5.2.3. The use of Models in Biology Lessons

Even though modelling is identified as one of the core scientific practices, activities such as creating models or formulating hypotheses using them are rare in biology education. Biology teachers mostly use models for illustrative purposes, thereby neglecting the scientific practice behind these models that is important to understand their use in science and learn about the nature of science (Windschitl et al., 2008). It therefore is unsurprising that students' understanding of biological models (e.g., Grünkorn et al., 2014) and biological concept-process models specifically (Jansen et al., 2019) has been shown to be limited.

Lack of experience with scientific modelling may be a major reason for biology teachers in limiting them in educating students about this scientific practice (Windschitl et al., 2008). Fortus et al. (2016) found that the STEM curriculum and textbooks that they analysed provide no explanation of what a model is, what the process of modelling entails, or why meta-modelling knowledge is important when working with models. This can make it difficult for teachers to explicitly teach students the required meta-modelling knowledge to develop an understanding of models as they are used in science. As described by Duschl, Schweingruber and Shouse (2007), teachers and students need support to understand and apply the core ideas behind scientific modelling.

Much research has been done on teaching and learning approaches that address reasoning with models in science (e.g., Gilbert & Justi, 2016; Heijnes, van Joolingen, & Leenaars, 2018; Krell et al., 2012; Quillin & Thomas, 2015; Schwarz & White, 2005; Treagust et al., 2002). One approach that stands out is the use of *drawing* as a means to learn about modelling (e.g., Ainsworth et al., 2011; Gijlers et al., 2013; Quillin & Thomas, 2015).

5.2.4. Drawing to learn about Models in Science

Drawing has been integral to the practice of science, since it is often used in the generation of hypotheses, the design of experiments, the visualization of data and the communication of findings to others (Ainsworth et al., 2011; Schwarz et al., 2009). Quillin and Thomas (2015) combined best practices from literature and suggest drawing-based interventions that can be used to stimulate model-based reasoning. They define multiple goals for drawing. Drawing models can help students to construct their own knowledge by connecting concepts, processing data, solving problems, or designing and interpreting experiments. Teachers can use students' drawings as a diagnostic tool to elicit their mental models about a certain phenomenon, such as their conception of how genes are related to evolution, including potential misconceptions (Dauer, Momsen, Speth, Makohon-Moore, & Long, 2013).

When it comes to drawing and modelling, Quillin and Thomas (2015) point out a difference between novice and expert learners, where novice learners tend to view models as static summaries of reality, and expert learners view models as flexible thinking tools. Their study suggests interventions that can both serve to develop teaching activities that fit the biological topic at hand, and as testable hypotheses for biology education researchers. The interventions focus on 1) improving affect towards drawing models, 2) improving *visual literacy,* and 3) improving visual model-based reasoning via drawing.

5.2.5. Visual Literacy

Visual literacy can be defined as the skill to read and write visual or symbolic language, including the ability to translate verbal models to visual models (e.g., Schwamborn, Mayer, Thillmann, Leopold, & Leutner, 2010; Stern, Aprea, & Ebner, 2003; Van Meter, Aleksic, Schwartz, & Garner, 2006), visual models to visual models (e.g., Johnstone, 1991; Novick & Catley, 2007) and visual models to verbal models (e.g., Schönborn & Anderson, 2010). These different components are illustrated in Figure 5.2, using the chromosome as an example.

Since biological processes manifest themselves on multiple levels of biological organization, multiple visualizations – or models – showing the same biological process can exist for each level of biological organization. Figure 5.2 shows that applying horizontal translation to a model of a chromosome means that the transition occurs on the same level of biological organization. In this case two different representations of a chromosome at the cellular level are visualized. Vertical translation means translating a drawing from one level of biological organization to another. Knippels and Waarlo (2018) support this idea by indicating that in order to promote coherent conceptual understanding of biological phenomena, yo-yo thinking – which links to horizontal and vertical translation in and between the levels of biological organization – should be part of the curriculum. In Figure 5.2 a condensed chromosome is visualized at the cellular level, versus a segment of DNA as a part of a chromosome at the molecular level.

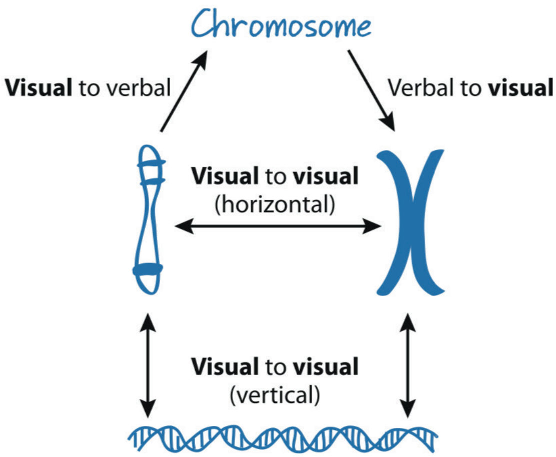


Figure 5.2. A visualization of the different components that are part of visual literacy (reprinted with permission from Quillin & Thomas, 2015).

Considering biological concept-process models specifically, the visual language in a model usually consists of representations of model entities (e.g., an organ) and arrows that represent relations between the entities and/or dynamic processes. To find out what learning strategies are beneficial, specifically when reasoning with biological concept-process models, Kragten, Admiraal and Rijlaarsdam (2015) observed differences in adopted learning strategies between more and less successful students when studying these models. They found a difference in strategy between successful learners and non-successful learners in which 80% of student-level variance in comprehension score was explained by the number of process arrows a student gave meaning to. These results support the idea that for understanding and creating concept-process models it is essential to understand the visual language in which they are represented. This emphasizes the need for an instrument assessing students' meta-modelling knowledge, in order to find out the effect of applying interventions such as the ones described by Quillin and Thomas (2015) on students' meta-modelling knowledge considering concept-process models.

5.2.6. Assessing Students' meta-modelling Knowledge

Several frameworks aiming at assessing students' meta-modelling knowledge have been developed. Most focus on aspects that are important for understanding models and their use in science, such as the nature and purpose of models (e.g., Crawford & Cullin, 2005; Grosslight et al., 1991; Justi & Gilbert, 2003) or handling models, which involves more practical skills (e.g., Louca et al., 2011; Schwarz et al., 2009). Upmeier zu Belzen and Krüger (2010) combined the frameworks from Crawford and Cullin (2005), Grosslight et al. (1991) and Justi and Gilbert (2003) and created an analytical framework for assessing and investigating students' understanding of models and their use in science. The five main aspects of this framework are *nature of models, multiple models, purpose of models, testing models* and *changing models.* For each of these aspects, several levels of understanding were formulated, ranging from an initial level to an expert level (Level 3). Grünkorn et al. (2014) evaluated this framework in the context of biological models, after which Jansen et al. (2019) elaborated some of the aspects to include biological concept-process models (Table 5.1). Since the framework by Jansen et al. (2019) includes categories that are specific for concept-process models, this version of the framework will be used in the current study to assess students' expressed meta-modelling knowledge of biological concept-process models.

The aspects *nature of models* and *multiple models* reflect the fact that models can describe and simplify phenomena, whereas the other three aspects (*testing models, changing models* and *purpose of models*) reflect the way models are *used* in science: to communicate ideas, make predictions about future events and test hypotheses (Grünkorn et al., 2014). In general, Level 1 focuses only on the model as a close representation of the original (*model object,* Mahr, 2009). When the model is seen as a medium that has been created based on the original, Level 2 has been reached (*model of something,* Mahr, 2009). Level 3 shows a comprehension of the use of models in science, in which models are used to test hypotheses and draw conclusions about the original (*model for something,* Mahr, 2009). In contrast to Level 1-3, the initial level that has been formulated for the aspects *testing models, changing* models and *multiple models* cannot be seen as a valid way of reasoning as it shows a lack of basic understanding of these aspects, rejecting the fact that models can be tested or changed, or that multiple models of the same process may exist (Grünkorn et al., 2014). It was empirically shown by Krell et al. (2012), that the Levels 1, 2 and 3 as mentioned in Table 5.1 reflect an increasing degree of difficulty.

This study focusses on the first three aspects within the framework: *nature of models, multiple models* and *purpose of models*. Looking at concept-process models specifically, in relation to the aspect *nature of models* students can argue

Table 5.1. Framework to assess students' meta-modelling knowledge of biological concept-process models. The left column shows the five aspects that are important when reasoning with biological models. For each of these aspects up to four levels of understanding have been defined, ranging from an initial level of understanding to an expert level of understanding. Categories in bold have specifically been added to assess students' meta-modelling knowledge of biological concept-process models (Jansen et al., 2019).

	Initial level	Level 1	Level 2	Level 3
Nature of models		- Model as copy - Model with great similarity - Model represents a (non-) subjective conception of the original **- Displays a process, its components and how they are related**	- Parts of the model are a copy - Model as a possible variant - Model as focused representation	- Model as hypothetical representation
Purpose of models		- Model for showing the facts **- Model for showing events**	- Model to identify relationships	- Model to examine abstract/ concrete ideas
Multiple models	- All models are the same - Various models of different originals - Only one final and correct model	- Different model object properties	- Focus on different aspects	- Different assumptions
Testing models	- No testing of models **- Perceiving schoolbooks or their authors as authorities providing absolute truth**	- Testing of material - Testing of basic requirements	- Comparison between original and model - Comparison and matching of original and model	- Testing hypotheses - Testing of hypotheses with research designs
Changing models	- No reason for alterations - Alteration of how different originals are represented	- Alterations to improve the model object - Alterations when there are errors in the model object - Alterations when basic requirements are not met	- Alterations when model does not match the original - Alterations due to new findings about the original	- Alterations due to findings from model experiments **- Alterations when the focus of the model shifts to a different aspect of the process**

that the model displays a process, different components and how they are related, without linking the model to the original (Level 1); be aware of the fact that models are often simplifications, fitted to the questioners' need and prior knowledge, and therefore only focus on a part of the process (Level 2); or know that models can be used to highlight a certain hypothesis relating to the depicted process (Level 3) (see Table 5.1). The aspect *multiple models* addresses the fact that multiple models can be used to represent the same original.

No single concept-process model can show every detail of a process or explain all existing hypotheses and ideas. This leads to a variety of models addressing the same process, but varying in aesthetic choices (Level 1), showing a difference in focus (Level 2), or explaining different ideas (Level 3). For the aspect *purpose of models*, concept-process models can be used to show various steps in a biological process (Level 1), they can be used to explain relationships between a process and other biological concepts or processes (Level 2), or they can be used to examine ideas and formulate hypotheses (Level 3).

The different levels of modelling as identified in the framework do not imply that lower levels are wrong. Instead, Level 1, Level 2 and Level 3 show different ways models can be seen or used, where the higher levels reflect the ways in which models are used in the course of scientific reasoning.

5.2.7. Reasoning with existing biological concept-process Models

Improving knowledge on models and the use of models in science is important when stimulating students' meta-modelling knowledge (Grosslight et al., 1991). Educating students about the aspects and levels as described in Table 5.1 can aid in improving this knowledge on models and their use in science. Considering concept-process models specifically, we previously looked into students' expressed meta-modelling knowledge considering the aspects *nature of models, multiple models* and *purpose of models* (Jansen, Knippels, & van Joolingen, 2021a). In this study we developed teaching and learning activities to familiarize students with the aspects *nature of models, purpose of models* and *multiple models*. Results showed that students were able to explain the meaning of the aspects *nature of models, purpose of models* and *multiple models*. Also, they were able to apply the aspects to different biological concept-process models and used language that was more in line with the descriptions of the aspects as shown in Table 5.1. However, their expressed levels of meta-modelling knowledge considering the three aspects of focus (Table 5.1) did not improve significantly. It has to be noted though that these studies focused on students' reasoning with existing biological concept-process models, without incorporating model-based practices. Since it often has been argued that model-based practices are necessary to stimulate reasoning with models in science (e.g., Gilbert & Justi, 2016; Heijnes et al., 2018; Krell et al., 2012; Quillin & Thomas, 2015; Schwarz & White, 2005; Treagust et al., 2002), the current study focusses on the effect of model-based practices on students' expressed levels of meta-modelling knowledge related to biological concept-process models.

5.2.8. Aim of the Study

We aim to design teaching and learning activities that foster students' visual literacy and model-based reasoning with biological concept-process models, and to assess the influence of these activities on students' expressed meta-modelling knowledge. A team of teachers and researchers is formed to make pedagogical and didactical choices when integrating the suggested activities in the design of two lessons. This study follows a two-step model, where key activities are developed, after which the influence of these activities on students' model-based reasoning and expressed meta-modelling knowledge is assessed. This leads to the following two research questions:

> *5.1: What Key activities are developed by the LS-team to foster students' visual literacy and engage students in modelling processes with biological concept-process models when following the suggested interventions by Quillin and Thomas (2015)?*
>
> *5.2: What is the influence of the developed key activities on students' reasoning with biological concept-process models?*

5.3. Method

In this study we first develop model-based reasoning activities that focus on biological concept-process models. Second, we assess the effect of these activities on students' expressed meta-modelling knowledge.

5.3.1. Participants

In total 61 Dutch eleventh-grade pre-university level students (16 – 18 years old, 32 male, 29 female) participated in our study. Three biology teachers (18, 30 and 9 years of experience, one male, two female) from the same secondary school formed the team that designed the key activities, together with the first author who is also a secondary school biology teacher with eight years of experience. The second and third author joined this team regularly as process facilitator and – during the execution of the research lessons – as observer. All lessons were taught in the same classroom as where the students' biology lessons usually take place. This study was performed in accordance with the ethical guidelines of the Science Faculty of Utrecht University. Informed consent was obtained from students and their parents as well as from the participants in the LS-team.

5.3.2. Lesson design

In this study we use Lesson Study (LS) to design and study key activities to foster students' visual literacy and engage students in modelling processes with biological concept-process models. LS is a Japanese model in which a team of teachers (LS-team) works together to design, teach, observe and evaluate lessons that are based on both literature and practice (Fernandez & Yoshida, 2004). This strong involvement of teachers allows them to integrate their experience and expertise into the design.

In this study, the LS-team designed two lessons following the LS-cycle as described by de Vries et al. (2016). Following the LS variant as described by Jansen et al. (2021a), both the teachers and researchers had a distinct role in the LS-team. The teachers made the pedagogical and didactical choices considering the design of the lesson, and the researchers checked whether these choices were backed up by literature. The lessons followed the suggested interventions by Quillin and Thomas (2015). The first lesson focused on stimulating students' visual literacy, the second on stimulating model-based reasoning via drawing. Four two-hour meetings to design the two 90-minute lessons took place. All meetings were audio recorded. In these meetings the LS-team 1) was presented with literature on *visual literacy,* model-based reasoning and meta-modelling knowledge, after which the research questions for this study were discussed, 2) decided on which curriculum topic the LS-team wanted to focus and what kind of concept-process models they thought were important to use, 3) focused on formulating key activities and 4) decided on three case-students that would be observed in detail during each lesson and on predicting the learning behaviour of these three students.

The expertise of the teachers in the LS-team and their knowledge about the students were used for the selection of the case-students. Teachers were asked to define for every student if they thought the student would find the developed key activities to be hard, medium, or easy to complete. They were also asked which students would be explicit in their arguments, making it easier to follow their way of reasoning during the lessons. The case-students that were selected by the teachers were observed in detail during the lessons. For each of these students an observation scheme was set up by the LS-team, describing predicted behaviour of this student during each phase of both lessons. In case one of the selected students would not attend the lesson, a back-up case-student for each of the case-students was chosen and an observation scheme for these back-up case-students was formulated. Table 5.2 shows an example of part of the observation scheme.

Table 5.2. Part of the observation scheme that was used during the research lesson designed by the LS team (translated from Dutch). The script includes the teacher activity and expected student behaviour for each phase of the lesson as formulated by the LS team. In this table only part of one of the phases (assignment 1) is depicted.

Phase of the lesson	Teacher activity	Expected student response	Notes
Assignment 1 (5 min) **13.30h - 13.35h**	**1.30 PM:** Start the assignment: *The first assignment focuses on understanding the meaning of symbols and colors that are used in a model. Use your work sheet to create a legend for this model about the regulation of body temperature. You will have to address the meaning of arrows, blocks, and the different colors that are being used in this model. You will have five minutes to complete the assignment.*	Student A: starts working on the assignment the moment the teacher finishes explaining what to do. Understands what to do and doesn't need any time to think before answering the questions Student B: Takes the assignment serious, but has to think for a while before answering the questions Student C: Finds it difficult to answer the questions and does not finish the assignment	

One of the teachers from the LS-team taught both lessons. The other members of the LS-team each observed one of the case-students during both lessons, following the observation scheme. Discussions between students taking place in the student groups that contained case-students were audio recorded separately and all student work was collected. Following the steps as described by de Vries et al. (2016), after teaching a lesson for the first time to one of the biology classes the lesson was discussed within the LS-team and improvements were formulated. The adjusted lesson was taught one week later by the same teacher in a different biology class of the same level at the same school. Observations and recordings were made in the same way as for the first lesson execution. The second lesson, addressing model-based reasoning, was taught one week after the first lesson. In a similar way as for the first lesson, this second lesson was also discussed, adjusted and taught a second time.

5.3.3. Data sources

5.3.3.1. Interviews – Case Students

All six case-students, three from each of the two biology classes, were interviewed after both lessons using the semi-structured interview scheme as described by de Vries et al. (2016). These questions related to the way students experienced the lesson. They were asked what they liked about each lesson; what element of the lessons motivated them the most, what they learned from the lessons; which element of the lessons they experienced as 'easy'; which element they experienced as 'difficult'; and what they would change about the lessons if these lessons would be taught again to a different class. Next to these questions, students were asked for the first lesson to formulate their own definition of the term visual literacy. For the second lesson students were asked to define what they thought was important when creating their own model. Interviews were audio-recorded and lasted 5-10 minutes.

5.3.3.2. Interviews – Teachers

To find out how teachers experienced developing the lessons, all three teachers who participated in the LS-team were interviewed after the completion of the LS-cycle using a semi-structured interview scheme. Interviews were audio-recorded and lasted approximately 40 minutes. The teachers were asked to reflect on their experience with the LS-cycle in which they developed the lessons to stimulate visual literacy and model-based reasoning. The interview questions were related to the expectations the teachers had before starting the intervention and to what extent these expectations were met; what they learned from participating in a LS-cycle; the extent to which they applied what they learned to other lessons or their teaching approach; and whether they expected to keep on using what they had learned on the long term.

5.3.3.3. Pre-test and Post-test (developed lessons)

To investigate what the effect of the two lessons was on students' understanding of visual literacy (lesson 1) and model-based reasoning (lesson 2), an online test was developed containing seven open-ended questions. The questions were based on the interventions proposed by Quillin and Thomas (2015) to stimulate visual literacy and model-based reasoning via drawing. The following question is an example of one of the questions in this test. It addresses the use of symbols in models, which is part of visual literacy:

> *Models often contain symbols, such as arrows. In the picture below you see three examples of models containing arrows. Describe what, according to you, is the general meaning or definition of an arrow in a model of a biological process. Note: this question is about the use of arrows in models in general, not about the meaning of a specific arrow in one of the models pictured below [question is followed by three biological concept-process models].*

Other questions related to students being able to filter out important aspects from a text to be represented in a model, giving reasons for the existence of multiple models of the same process, formulating the main message from a given model, and explaining why the elements in a given model should be arranged in a specific order. A list containing all translated questions can be found in Appendix D.

5.3.3.4. Pre-test and Post-test (full academic Year)

The current study took place in the same academic year as the study described in Jansen et al. (2021a) in which the same students participated. This allowed us to look at the effect of the combination of these two studies on students' expressed meta-modelling knowledge. To find out the effect of this combination of studies, we used an online test that was previously developed based on work from Krell (2019) to assess students' level of understanding for biological concept-process models (van Montfort, 2019; Jansen et al., 2021c). This test was evaluated and used previously to measure the average level of students' understanding in the Netherlands for biological concept-process models (van Montfort, 2019; Jansen et al., 2021c). The test contains statements related to six different biological concept-process models. For each of these models a statement on Level 1, Level 2 and Level 3 is formulated for the aspects *nature of models* and *multiple models* (see Table 5.1 for a description of the levels). The test only focuses on two aspects, since including all aspects would generate too many questions for students to answer. The choice for the aspects *nature of models* and *multiple models* was made since these aspects reflect on the way students reason with models that are presented to them, resembling both the way students are presented with models during the test and the way students encounter scientific information in daily life. In this test, students mention whether they agree or do not agree with the given statements. Students' agreement with the statements is interpreted as their expressed meta-modelling knowledge for the aspects *nature of models* and *multiple models* (Krell, 2019; Jansen et al., 2021c). Figure 5.3 shows an example of one of the tasks in this test, containing a model showing the process of bioaccumulation. In this case the task reflects on the aspect *multiple models*. Students filled out the test as a pre-test at the beginning of the academic year, and as a post-test at the end of the academic year. A link to the complete online test can be found in Appendix C.

5.3.4. Data Analysis

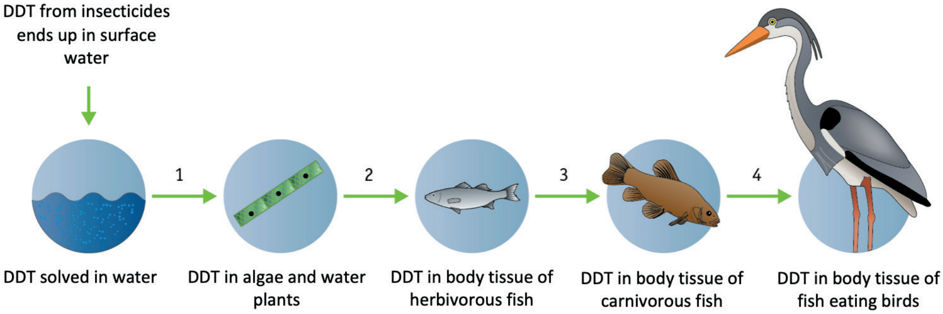
Data analysis will be described in line with the two research questions.

> *RQ1: what key activities are developed by the design team to foster students' visual literacy and engage students in modelling processes with biological concept-process models when following the suggested interventions by Quillin and Thomas (2015)?*

To answer the first research question, the interventions as described by Quillin and Thomas (2015) to improve students' visual literacy and model-based reasoning via drawing were combined with pedagogical approaches and teachers' knowledge and experience to design two lessons. The role of the teachers and the researchers in the LS-team was leading in the development of these lessons. The first lesson

Example item

Frank is doing research on water contamination. He discovered that in areas where the water is contaminated with DDT (a chemical insecticide), many fish eating birds are dying. He creates the following model.



Maud is also doing research on water contamination with DDT. She creates a model of the same process. When can the model that Maud creates be defined as a different model than Franks' model?

Answer for each of the statements if you agree (yes) or disagree (no)

When Maud creates a model for an area that is contaminated with DDT, but where different species of fish are living.

[ ] Yes

[ ] No

When Maud creates a model where the focus is more on step 2 of the process.

[ ] Yes

[ ] No

When Maud has different thoughts on how DDT ends up in fish eating birds and creates a model reflecting these thoughts.

[ ] Yes

[ ] No

Figure 5.3. An example of one of the tasks in the test, showing a model of bioaccumulation. The aspect of focus in this example is multiple models and each statement represents a level of understanding in the order: Level 1, Level 2 and Level 3 (the order of the levels differed between different tasks). The model present in this task is reprinted and translated with permission from Bos et al. (2012).

focused on improving students' visual literacy, while the second lesson addressed model-based reasoning. Recordings from the LS-meetings (six meetings, two hours per meeting) in which the lesson was designed, were used to distil the pedagogical and didactical choices made for each of the key activities.

> *RQ2: what is the influence of the developed key activities on students' model-based reasoning and expressed meta-modelling knowledge considering biological concept-process models*?

To find out the effect of the two lessons addressing visual literacy and model-based reasoning, both data from the open-ended pre- and post-test covering these two lessons and data that was gathered during the two lessons (student products, student interviews and audio recordings from the lessons) within this LS-cycle was analysed.

Considering the pre- and post-test, only students who participated in both lessons and filled out the pre- and post-test were incorporated in the data analysis (n=37). Table 5.3 shows the scoring mechanism for each of the questions in this test. A paired t-test was used for Questions 1, 2, 3 and 6 to compare students' results from the pre-and post-test. For questions where the answer was scored as 'correct' or 'wrong', a related samples McNemar test was used (Questions 4, 5 and 7).

Audio recordings from the lessons, records from the LS-meetings and student interviews were scanned for utterances related to what students thought was the most enjoyable part of the lesson, the most motivating part of the lesson, which aspects they learned from the lesson, how they defined visual literacy, and what aspects students considered to be important when creating their own model. These data were used to complement the results from the open-ended pre- and post-test.

The results from the pre- and post-test covering the full academic year were analysed to determine the effect of the combination of the current study and the lessons as described in Jansen et al. (2021a) on students' expressed meta-modelling knowledge for the aspects *nature of models* and *multiple models*. Only students who filled out both the pre- and post-test were included in the data analysis (n=45). The results of the test showed for six different biological concept-process models whether students agreed with statements on Level 1, Level 2 and Level 3 for the aspects *nature of models* and *multiple models*. The agreement score for all levels and aspects was determined separately for each student. Using a paired t-test we calculated whether the difference between the number of times a student agreed with statements on a certain level in the pre-test and post-test could be considered significant.

Table 5.3. Questions from the pre- and post-test and the scoring mechanism for each of these questions.

Question	Type of question	Scoring mechanism
1	Giving meaning to the use of arrows in models	0 = wrong answer 1= describing a relationship between aspects 2 = using the word 'dynamic' - or a synonym of this word - when describing the relationship between aspects
2	Defining important aspects from a text that should be incorporated in a model of a biological process	0 = no aspects defined 1 = one or two aspects defined 2 = three or four aspects defined
3	Formulating a reason for the existence of multiple models of the same biological process	Scored using the Levels as described in Table 5.1 for the aspect *multiple models* (Level 1, 2 and 3)
4	Defining whether in the end there will be one ultimate model of a biological process	0 = yes (wrong) 1 = no (correct)
5	Formulating the key message of a model	0 = wrong 1 = correct
6	Explaining the role of a 'sensor', a 'control center' and an 'effector' in a model that describes homeostasis	0 = no aspects correctly explained 1 = one aspect correctly explained 2= two aspects correctly explained 3 = three aspects correctly explained
7	Choosing the correct order in which aspects in a model should be related to each other (3 options are given)	0 = wrong option 1 = correct option

5.4. Results

The results will be described in line with the two research questions.

5.4.1. RQ5.1: Key Activities

The teachers in the LS-team made pedagogical and didactical choices for the lessons, after which the researchers checked whether these choices are backed up by literature. Below the choices of the teachers, consisting of both general pedagogical choices and specific pedagogical choices for each of the lessons, are listed. Each choice is accompanied by literature references that were brought in by the researchers in the LS-team.

5.4.1.1. General pedagogical Choices (both Lessons)

In order to get students engaged with the theory about visual literacy and important aspects of model-based reasoning, the LS-team agreed to adopt an inquiry-based approach in which students work with these aspects themselves first, before the theory was explained to them. As shown by Gormally, Brickman, Hallar and Armstrong (2009), inquiry-based learning stimulates scientific reasoning and helps students to gain confidence in their scientific abilities. Since model-based reasoning is known to be a scientific metacognitive skill, inquiry-based learning could aid in developing this ability.

The LS-team decided that students had to talk out loud during the lesson while working together in groups. Tanner (2009) showed that talking out loud can improve metacognition, since it promotes student thinking about aspects that they understand and aspects they do not understand. According to Smith et al. (2009), working in groups can aid in learning and improve student performance in general. Talking out loud also allows the observers during the execution of the lesson to better follow students' thinking processes.

5.4.1.2. Specific pedagogical Choices for Lesson 1 – Visual Literacy

As argued by Quillin and Thomas (2015), visual literacy needs to be addressed before reasoning with biological models can be stimulated. To teach students how to select important information and connect important concepts within a model, without being distracted by surface features, practice is necessary (Dauer et al., 2013; Harrison & Treagust, 2000; Van Meter et al., 2006). The interventions proposed by Quillin and Thomas (2015) to stimulate visual literacy include drawing models in which students practice to pick out important information and connect important concepts. They argue that in order to enhance students' visual literacy, symbols in models need to be explicitly defined in class and students need to practice both with the drawing medium - such as paper and pencil - and with horizontal and vertical translations (Figure 5.2). The LS-team followed the proposed interventions by Quillin and Thomas (2015) to design key activities that stimulate visual literacy using biological concept-process models.

5.4.1.3. Specific pedagogical Choices for Lesson 2 – fostering model-based Reasoning

Literature has shown that there are differences between novice and expert learners regarding the use of models in various STEM disciplines (Harrison & Treagust, 2000; Hmelo-Silver, Marathe, & Liu, 2007; Schwarz et al., 2009), where novice learners for example show to have more trouble understanding causal behaviours and functions (Hmelo-Silver et al., 2007). Quillin and Thomas (2015) propose interventions to improve visual model-based reasoning via drawing, where students practice with the use of models in biology. In the proposed interventions students need to draw a model, use their model to answer questions, evaluate their model and revise their model. The LS-team decided to base the design of the learning and teaching material on these interventions and see whether these activities contributed to a higher level of expressed meta-modelling knowledge as shown in Table 5.1.

5.4.1.4. Design Lesson 1: Visual Literacy (90 min.)

To avoid students linking the theory from this lesson to one specific biological topic, the LS-team decided to use models about various biological topics. However, in order to make sure students were familiar with these biological topics, only models relating to the chapter they were working on at that time and the previous chapter they worked on were chosen as suitable models for this lesson. The lesson consisted of five key activities, all based on the interventions as proposed by Quillin and Thomas (2015) to promote visual literacy. The proposed interventions as formulated by Quillin and Thomas (2015) are mentioned below, followed by the corresponding key activity as formulated by the LS-team. Quotes from the LS-team are added to give more insight into why the teachers decided on these activities.

Intervention: explicitly define the symbols used in class, both "generic" symbols (such as axes in a graph) and subdiscipline-specific symbols (such as branches in a phylogenetic tree)

Key activity 1.1: Students define the meaning of arrows, blocks and colours in a model of how the human body maintains its temperature.

R: So maybe it would be good to select a subject for our first activity?

[...]

T1: The one about the body temperature seems suitable. It contains arrows and blocks [referring to visual literacy as defined by Quillin and Thomas (2015)].

T2: Yes, the chapter in the book also starts with that one, so it can be a good starting point.

Intervention: give students opportunities to practice translating text to drawings

Key activity 1.2: students translate a text about how the body maintains blood pressure into a drawing.

T2: For this activity we can do something with hormones. We used to have an activity where they had to design a poster about a process, so we kind of already had this type of assignment in previous years. We can now let them draw a model instead.

Intervention: give students opportunities to practice translating "horizontally" from one drawing to another at the same scale

Key activity 1.3: Students translate a model showing a drawing of cells undergoing mitosis (cell division) into a table that defines all steps of this process and includes the amount of DNA present in the cells at each step.

T3: It is sometimes hard for students to understand, when you talk about cell division, that DNA replicates but the cell stays 2n [diploid]. So when you attach a letter to the amount of DNA in the cell, it makes things more clear. This can be read from a drawing and then put into a scheme.

T2: So then we kind of have a translation from a model into a scheme

Intervention: give students opportunities to practice translating "vertically" from a drawing at one scale to a drawing at another scale

Key activity 1.4: students draw a model of a plant cell, after which they zoom in on the cell membrane of this cell.

> *T2: I think the example from the article [referring to Quillin and Thomas, 2015] will be suitable for our students. Draw a model of a plant cell with at least five organelles. Then zoom in on the cell membrane, right? We discussed this topic in class previously, so that should work.*

Intervention: give students opportunities to practice translating drawings to text

Key activity 1.5: students formulate what they think is the main message of a graph showing the enzyme activity for different enzymes at different temperatures.

> *T3: It would be good to use a graph in one of the assignments. I feel like students sometimes have trouble reading graphs.*
>
> *R: But it has to be a model of a process, because that is the focus of our research*
>
> *T3: You also have graphs showing a process, like the ones where you have to find out which enzymes work inside a human body.*
>
> *T2: Right, the ones where you can see the optimum temperatures for different enzymes.*

After completing these key activities, students compared their answers in groups of four. Together they discussed the differences between their answers and decided for each activity which of the answers in the group they thought best answered the question. Several groups shared their ideas plenary, after which the teacher provided the correct answers and explained why these answers were correct. Students then received a list with the activities within this lesson, which they were asked to rank on a scale from 1-5, with 1 very difficult and 5 very easy. Table 5.4 shows how the key activities as formulated by the LS-team correspond to the interventions as proposed by Quillin and Thomas (2015).

5.4.1.5. Design Lesson 2: model-based Reasoning (90 min.)

The topic of this lesson was in line with the chapter that was scheduled for this time in the schoolyear: homeostasis, the process of maintaining steady conditions in living organisms. In this chapter many topics are discussed, from which the LS-team chose 'blood-glucose regulation', the way the human body maintains the amount of glucose in the blood around a set level. The lesson had to be the first lesson that the students received on this topic, since students had to engage in modelling practices without having an existing model of this process in mind. The following conversation shows how the LS-team decided on the topic blood-glucose regulation:

T2: I think the topic of the model should be something that they would be working on anyway in class

T1: Yes, also because this lesson will have to replace some other regular lessons. Otherwise we will not have enough time to finish our curriculum. But it has to be a model that can be changed, because that is one of the activities in the article.

T2: Yes, but you can always introduce a disease.

T1: So, it has to be something where we can introduce an external influence that is easily understood by the students, without a lot of complicated extra theory.

T3: So, something like insulin?

T1: Yes, and then you can talk about diabetes? That's usually also discussed in class.

T2: So, I think, we could give the students a story about glucose regulation and let them create a model using this text?

Table 5.4. Summary of the interventions as described by Quillin & Thomas (2015) to stimulate visual literacy and the corresponding key activities as formulated by the LS-team for research lesson 1.

Intervention	Key activity
1. Explicitly define the symbols used in class, both "generic" symbols and subdiscipline-specific symbols	Students define the meaning of the arrows, blocks and various colors that are used in a model about the maintenance of human body temperature
2. Give students opportunities to practice translating text to drawings	Students translate a text about how the body maintains blood pressure into a drawing
3. Give students opportunities to practice translating "horizontally" from one drawing to another at the same scale	Students translate a model showing a drawing of cells undergoing mitosis into a table
4. Give students opportunities to practice translating "vertically" from a drawing at one scale to a drawing at another scale	Students draw a model of a plant cell and draw a second model of the same cell zooming in on the cell membrane
5. Give students opportunities to practice translating drawings to text	Students formulate the main message of a graph

The eight activities in the lesson were all based on the interventions as proposed by Quillin and Thomas (2015) to stimulate model-based reasoning and related to the topic blood-glucose regulation. Each proposed intervention is mentioned below, followed by the key activities as formulated by the LS-team. Quotes from the LS-team are added to give more insight in the reasons for choosing these activities.

Intervention: explicitly point out the difference between surface features and structural features (the underlying relationships, processes, functions, and principles in the models) and explicitly walk through the process of creating a model for students before asking them to make their own

Key activity 2.1: To guide students through the modelling process, they first answer questions about blood-glucose regulation, using a text that explains this process. They also define important aspects from this text that have to be incorporated into a model about blood-glucose regulation. Students then draw a model of the process. The teacher explicitly points out the difference between surface features (the appearance of the organs in the model) and structural features (the underlying relationships and processes in the model), stressing that the students should mainly focus on the structural features when drawing their model.

> *T1: So, when we do that, it [the proposed intervention] says that we have to explain what are important features and what are not important features when drawing a model. So what are the things we think are important to draw and what things are not so important to draw? So what are surface features and what are, as they call them here, structural features? So what are things they do not have to draw in detail?*
>
> *T3: They do not really have to draw the organs. They can, but it should not be the thing that they spend a lot of time on.*
>
> *T2: Yes, what you usually see is that the topic of the model is drawn with a lot of detail, to really emphasize or focus on that issue.*
>
> *T3: But in this specific example we need to know what is a more superficial feature and what is a structural feature.*
>
> *T1: So, I guess the way the level is maintained is what they should focus on. What the organs look like for example is not important.*

Intervention: prompt students to check the quality of their models to ensure that they include all the essential elements in an accurate way

Key activity 2.2: students use a rubric, created by the LS-team (supplementary Table 5.1), to check whether all essential elements are included in their model in an accurate way.

> *T2: I think the suggestion in the article of using rubrics is quite nice. But then we have to create a rubric ourselves.*
>
> *T3: Yes, we can do something like, to what extent is a specific element in the model present or correctly represented?*
>
> *T1: Yes, that is something that should be quite easy to create.*

Intervention: prompt students to check the quality of their models to ensure that they are including only what is relevant.

Key activity 2.3: students exchange their model with their neighbour who reviews every step in the model and gives feedback on the elements that are present in the model.

> *T1: We could just use the suggestion from the article, let students exchange models with their neighbour and ask for feedback.*

Intervention: prompt students to make improvements on their models based on their (or someone else's) evaluation of it.

Key activity 2.4: students put their model in a clear plastic folder, after which they draw the improvements to their model on this folder.

> *T3: I think using clear plastic folders is an ideal way to do this. Then you can easily separate the original drawing from the improvements.*
>
> *T4: Yes, and everything is already organized in a folder! Also, you always have students that would make changes using the same pen or pencil as the one they used to draw the initial model. If you would let them draw on the original drawing you would probably not be able to see what the adjustments are.*

Intervention: prompt students to use the models they create as tools to answer questions

Key activity 2.5: students use their model to explain what happens inside a human body when someone eats five donuts, containing a lot of sugar. Their answer has to include several elements from the model.

> *T2: Well, this one is quite easy. Just let them answer what happens when they eat a lot of sugar. Five donuts for example, just like T3 mentioned earlier.*
>
> *T4: Or what happens when you start one of those diets where you should not eat a lot of carbohydrates.*
>
> *T2: Then you just eat the hole in the centre of the donut! No, let's just go with the donut, that is a nice idea because maybe they can also link it to diabetes later on.*

Intervention: prompt students to add or change an element in their model as a tool for solving a problem.

Key activity 2.6: students adapt their model to show what happens when someone is diabetic. Students then explain what a possible treatment would be for someone who is diabetic and why that treatment is appropriate.

> *T2: Yes, so then we can do something with the hormone insulin.*
>
> *T3: But they have to be able to predict what happens, without us explaining it to them.*
>
> *T1: Well, we can just use the topic diabetes, there is a text about that in our textbook.*

Intervention: demonstrate the flexibility of models by showing and prompting alternate versions of the same model

Key activity 2.7: students compare their model with a different model of the same process and define differences between these two models.

> *T2: We can link this assignment to the key activity we used in our first lesson [the lesson that was designed earlier that year, as described in Jansen et al. (2021a)].*
>
> *T1: Yes, we can link what they are doing now to our previous research lesson. Because this is about the aspect multiple models. We can again let them compare models and define differences.*

Intervention: demonstrate the metacognitive value of models by asking what parts of the model the student is struggling with

Key activity 2.8: students end the lesson by ranking the activities within this lesson on a scale from 1-5, with 1 very difficult and 5 very easy.

> *T2: We could just ask them, what part did you struggle with?*
>
> *T4: Or let them make a list with all the activities, asking them what was easy for you and what did you struggle with?*
>
> *T3: Create a top three of difficult and easy activities?*
>
> *T4: Perhaps just let them rank the activities on a scale from 1-5?*
>
> *T2: Yes, I think it is better to not just focus on what they struggled with, but to also let them think of what already goes well.*

After completing these key activities, the teacher discussed the answers in class. Table 5.5 summarizes the interventions as described by Quillin and Thomas (2015) and the corresponding key activities as formulated by the LS-team.

5.4.1.6. Adaptations of the Lessons

After teaching each of the lessons for the first time in one of the biology classes, the lesson was discussed in the LS-team as prescribed by the LS-steps described by de Vries et al. (2016) and improvements were formulated. Since the lesson mostly went according to plan, only minor adjustments to the lesson were made:

Lesson 1 - Visual Literacy:

Students mentioned that they were not very familiar with the theory for every topic that was discussed in this lesson. Even though all topics were part of the curriculum, some topics had been discussed in class a longer time ago. Students were therefore struggling to remember the exact theory. The LS-team decided that in Lesson 1 the teacher shortly had to explain the theory behind the models that were used in this lesson, before students started the key activities.

Lesson 2 – Model-based Reasoning

Students needed more time to draw their model than the LS-team had previously planned for. Therefore, the LS-team decided to extend the drawing time in Lesson 2 by ten minutes. Since the other activities in lesson 2 took less time than planned for, the added ten minutes drawing time did not lead to an extended Lesson duration.

Table 5.5. Summary of the interventions as described by Quillin and Thomas (2015) to stimulate model-based reasoning and the corresponding key activities as formulated by the LS-team for research lesson 2.

Intervention	Key activity
1. Explicitly point out the difference between surface features and structural features (the underlying relationships, processes, functions, and principles in the models) and explicitly walk through the process of creating a model for students before asking them to make their own	Students translate a text about blood-glucose regulation into a model. Students are led through the modelling process first by answering questions and defining important aspects from the text that had to be part of the model. Before the students start drawing the model, the teacher explicitly points out the difference between surface features (the appearance of the organs in the model) and structural features (the underlying relationships and processes in the model), stressing that the students should mainly focus on the structural features when drawing their model
2. Prompt students to check the quality of their models to ensure that they include all the essential elements in an accurate way	Students use a matrix to check whether all essential elements are included in their model in an accurate way
3. Prompt students to check the quality of their models to ensure that they are including only what is relevant	Students exchanged their model with their neighbor who reviews every step in the model and gives feedback on the elements that are present in the model
4. Prompt students to make improvements on their models based on their (or someone else's) evaluation of it	Students put their model in a clear plastic folder, after which they draw the improvements to their model on this folder
5. Prompt students to use the models they create as tools to answer questions	Students use their model to answer given questions about what happens inside a human body when someone eats a lot of sugar and are instructed to include several elements from the model in their answer.
6. Prompt students to add or change an element in their model as a tool for solving a problem	Students adapt their model to show what happens when someone is diabetic. They also explain what a possible treatment could be for someone with this disease
7. Demonstrate the flexibility of models by showing and prompting alternate versions of the same model	Students compare their model with a different model of the same process and define differences between these two models.
8. Demonstrate the metacognitive value of models by asking what parts of the model the student is struggling with	Students rank the activities within this lesson on a scale from 1-5, with 1 very difficult and 5 very easy.

5.4.2. RQ5.2: Effect on Students' expressed meta-modelling Knowledge

5.4.2.1. Open ended Pre-test and Post-test

To find out what the effect of the lessons on visual literacy and model-based reasoning was on students expressed meta-modelling knowledge, results from the pre- and post-test covering these two lessons were compared. Data from student interviews with the case-students were used to substantiate these results. Table 5.6 presents for each question the type of test that was used to see whether differences between the pre- and post-test were significant, the results of these tests and the mean student score for the pre- and post-test.

Table 5.6. The mean student score for each of the questions and the results from the statistical tests. For questions 1, 2, 3 and 6 a paired t-test was used. For these questions the mean represents the mean student score in the pre- and the post-test. For questions where the answer was scored as 'correct' or 'wrong', a related samples McNemar test was used (questions 4, 5 and 7). For these questions the mean represents the percentage of students that provided a correct answer (*M* = mean, n=37).

Question	Type of test	*M* Pre-test	*M* Post-test	*t*	*p*
1	Paired t-test	0.84	1.22	3.60	<.01
2	Paired t-test	2.92	3.00	0.24	.67
3	Paired t-test	1.59	1.76	1.64	.11
4	Related samples McNemar	86%	95%		.25
5	Related samples McNemar	43%	49%		.63
6	Paired t-test	1.51	2.14	1.98	.06
7	Related samples McNemar	73%	95%		.04

Students scored higher in the post-test than in the pre-test for all questions, but only the results for Question 1 and Question 7 showed a significant difference between the pre-test and the post-test. More students were able to provide a general meaning of the use of arrows in models in the post-test (M=1.22, SD=0.42) than in the pre-test (M=0.84, SD=0.37) (Question 1). Also, more students chose the correct order in which different aspects should be related to each other in a model about homeostasis in the post-test than in the pre-test (p=.04) (Question 7).

5.4.2.2. Interviews Case-Students

The quotes in this section are meant to substantiate the results from the pre- and post-test. All six case-students mentioned that they enjoyed the two lessons, because they differ from their usual classes in both pedagogy (e.g., working in groups instead of individually) and content. They mention that they are used to talking about the biological topic, but not to creating their own model, as shown by the following student quote. The model that was created by the student is added as supplementary material.

> *Student CS_CC: I learned to find the important aspects from the text. We have often seen text fragments like the one we received, but we have never created a model from text before [..] I think by doing this we are looking at the theory in a different way and learn to see the effects of something better. I think what I mostly learned is that by creating a model from text, you will get an overview of the theory, showing you what causes what and things like that.*

One of the case students mentions that creating models from text can be a new learning strategy. The model that was created by the student is added as supplementary material.

> *Student CM_CA: Creating a model is interesting. I learned a new learning strategy [...]. It is pretty easy, so I will be using this strategy when learning for my exams. I would create a model myself first, before using the models from the textbook, because using those will be less effective for me.*

Students also mention that they developed a more critical perspective on existing models, as shown by this student quote when talking about a model containing blocks in different colours:

> *Student CS_CB: I learned that you need a critical perspective when looking at a model. And that you can get more information from a model then you thought when you were looking at it at first. When you look at this model for example, I wouldn't have payed attention to the different colours at first. But when you do, you can see the relationship between different processes better.*

5.4.2.3. Pre-test and Post-test full academic Year

To assess the effect of the combination of lessons as described in Jansen et al. (2021a) and the current study on students' level of understanding for the aspects *nature of models* and *multiple models*, results from the pre- and post-test covering the full academic year were compared. Students' agreement with statements was interpreted as their expressed meta-modelling knowledge. This means that when a student is able to reason with models on the three levels as described by the framework (Table 5.1), we expect that student to agree with all three statements in a task. The test focused on the aspects *nature of models* and *multiple models.* Figure 5.4 shows the results of this test, in which Figure 5.4a shows the effect considering the aspect *nature of models* and Figure 5.4b shows the effect considering the aspect *multiple models.* Per student the number of times a student agreed with statements on a certain level was counted. The average number of times a student agreed with a statement is shown as a percentage, called the 'average agreement score'.

Agreement scores went up for all levels in both aspects, except for the aspect *multiple models* Level 1, where the agreement scores went down. Using a paired t-test, we found the increase in agreement between the pre- and post-test for the aspect *nature of models* Level 2 ($p = .03$) and the aspect *multiple models* Level 2 ($p < .01$) to be significant. The test also showed the decrease in agreement for the aspect *multiple models* Level 1 to be significant ($p = .01$).

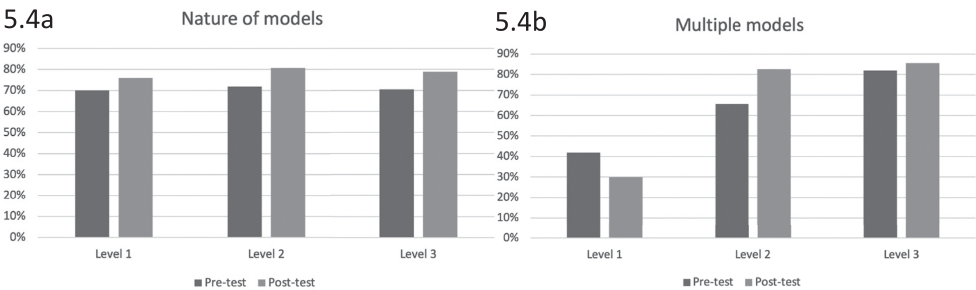


Figure 5.4. The results from the pre- and post-test covering the full academic year (n = 45). The percentage stands for the average agreement score. Figure 5.4a shows the average agreement score for the aspect *nature of models*, Figure 5.4b shows this score for the aspect *multiple models*.

5.5. Discussion

This study aimed to employ LS for designing teaching and learning activities focussing on concept-process models that follow the interventions as posed by Quillin and Thomas (2015) to stimulate visual literacy and model-based reasoning. Using an open-ended pre-and post-test and data from student interviews, we looked at the effect of these key activities on eleventh-grade pre-university level students. A pre- and post-test covering the full academic year was used to measure the effect of the combination of the key activities as described in our previous work (Jansen et al., 2021a) and the key activities as designed in this study, on students' expressed meta-modelling knowledge considering biological concept-process models.

To answer our first research question we used a Lesson Study approach, in which theory from literature and teacher experience are combined into a lesson design (Fernandez & Yoshida, 2004). In this design the interventions as proposed by Quillin and Thomas (2015) to stimulate visual literacy and model-based reasoning via drawing were used as a starting point, after which the LS-team developed key activities that focused on biological concept-process models specifically. The design consisted of two lessons in which the first lesson contained five key activities addressing the aspect *visual literacy*, and the second lesson contained eight key activities focusing on *model-based reasoning* via drawing. In the first lesson about visual literacy students focused on practicing with horizontal and vertical translation and understanding the visual language that is used in models. In the second lesson about model-based reasoning students focused on drawing a model and using, evaluating and revising this model.

The open-ended pre- and post-test combined with data from student interviews was used to evaluate the effect of the two lessons on students' visual literacy and model-based reasoning. Even though student results showed an increase for all aspects tested, most changes were minor. However, relating to visual literacy and model-based reasoning respectively, students showed great improvement in giving meaning to arrows and defining the correct order of aspects when creating a model (Table 5.6). Furthermore, data from student interviews shows that modelling activities were new to students and that they found the activities enjoyable and useful for learning. The activities also made students feel like they developed a more critical perspective on existing models. These results suggest that the activities as formulated have a positive impact on students' view on models and aid in developing a more scientific and critical approach when working with biological concept-process models. It has to be noted though that the students in this study only performed each of the activities once. It can be argued that, due to a difference in cognitive level between students, some students might need more practice than others. An instructor can, depending on the student, choose to scaffold modelling by first introducing formative exercises at lower-order cognitive levels before working up to assignments at higher-order levels (Quillin & Thomas, 2015).

To measure the effect of the combination of lessons as described in Jansen et al. (2021a) and the current study on students' level of understanding for the aspects *nature of models* and *multiple models*, results from the pre- and post-test covering the full academic year were compared. Results showed that students' agreement scores in the post-test were higher for all levels and both aspects, except for the aspect *multiple models* Level 1, where the agreement score was lower in the post-test than in the pre-test. A significant difference in agreement score between the pre-test and the post-test was observed for the aspect *nature of models* Level 2 and the aspect *multiple models* Level 1 and 2.

Students need to be able to reason on all three levels as described by Krell et al. (2012) and Jansen et al. (2021c), meaning that ideally they should agree with all levels in the pre- and post-test covering the full academic year for both the aspects *nature of models* and *multiple models*. Being aware of the different aspects of model-reasoning and their accompanying levels aids in developing a more scientific view on the use of models in science, where the lower levels within the framework focus on a more basic level of reasoning and the highest level reflects on the use of models in science (Grünkorn et al., 2014). The fact that an increase in agreement scores for multiple levels could be observed, suggests that the combination of the key activities, relating to different aspects of meta-modelling knowledge from our previous work (Jansen et al., 2021a) and the key activities relating to visual literacy and model-based reasoning as described in this article, leads to a more scientific view on the use of biological concept-process models for the aspects *nature of models* and *multiple models*. The decrease in agreement score for the aspect *multiple models* Level 1 could be explained by the way the interventions as described by Quillin and Thomas (2015) were implemented in the lesson design by the LS-team. The LS-team felt that a Level 1 type of reasoning for the aspect *multiple models*, showing a difference in appearance between models such as the use of colours or shapes, was very obvious and would therefore not be problematic for students. They agreed that the higher levels of reasoning, showing a difference in focus between models (Level 2) or a difference in assumptions (Level 3), were more important to address in class (key activities 1.3 and 1.4). Because of this choice, students might tend to agree more with Level 2 and Level 3 statements in the post-test, rejecting the Level 1 statements. However, further research should establish what influences the rejection of Level 1 statements for the aspect *multiple models*.

Since the pre- and post-test that covered the full academic year only focused on the aspects *nature of models* and *multiple models,* this study did not elaborate on the effect of the key activities on the level of students' reasoning for the other three aspects in the framework (*purpose of models, testing models* and *changing models)* (Grünkorn et al., 2014; Jansen et al., 2019). However, the suggested interventions as formulated by Quillin and Thomas (2015) show similarities with the aspects *testing models* and *changing models*. The proposed intervention where students have to check the quality of their model can be linked to the aspect *testing models*, while the intervention where students have to revise their model based on the evaluation of their model links to the aspect *changing models*. Further research should therefore address the effect of the key activities as formulated on students' expressed meta-modelling knowledge for the other aspects within the framework (Table 5.1).

It is important to note that the key activities in this study were developed using a LS-approach with a strong theoretical component. As described by Fernandez and Yoshida (2004), this means that the teachers from the LS-team are exposed to theory from literature, thereby learning things that other teachers might not be aware of. The following quote from one of the teachers in the LS-team shows that the theory that formed the basis of the lessons was new for this teacher:

> *T1: What we discussed in these lessons makes you realize that you should really pay attention to these concepts. These are things that I never really looked at myself, so that really was an eyeopener. The theory about visual literacy and reasoning with models was new to me.*

Interviews with the teachers in the LS-team show that being part of this team influences the teachers’ view on models in biology education. The following quote is from the teacher that taught the three lessons and shows the change in view on the way he addresses models in his teaching:

> *T2: I have seen that students need to learn how to interpret models. I used to think that this was something that would just come naturally to students. Because well, all information is present in the model. So, you would think it should be easy to deal with that information, right? This [being part of the LS-team] really changed my view on that. Also in other lessons, when I explain a model, I always describe what it is that we are looking at and what kind of choices the creator of the model has made when creating this model.*

The fact that the teachers’ view on models changed, can have an impact on the way the key activities are introduced and discussed in class. Therefore, this might also influence the effect these activities have on students’ level of reasoning with models. Future research should therefore focus on the effect of using a LS approach on the way the developed lesson is taught in class.

Our results give insight into how teachers develop activities using the suggested interventions as described by Quillin and Thomas (2015) using a LS approach. After introducing teachers to theory on concept-process models, visual literacy and model-based reasoning, our results suggest that teachers are able to combine their pedagogical and didactical knowledge (e.g., the effect of students working in groups or talking out loud on learning) with the theory that is presented to them and develop key activities that are in line with the interventions as proposed by Quillin and Thomas (2015). In the developed lessons students practiced with horizontal and vertical translation, understanding the visual language that is used in models, drawing a model themselves and using, evaluating and revising this model. The use of LS as a method to develop these lessons was a good choice for this study, since it allowed researchers to present the required theory to the teachers, and check whether pedagogical and didactical choices were in line with both the theory from literature on these subjects and with the suggested interventions as described by Quillin and Thomas (2015).

Results from the pre- and post-test covering a full academic year showed that the lessons on visual literacy and model-based reasoning as developed in this study contributed to a more scientific view on the use of concept-process models in science. LS as used in this study functioned as a bridge between theory and practice and showed how teachers not only use the presented theory for developing these two lessons, but extend this knowledge to other lessons and thereby implement the theory sustainably into their teaching. Since the creation and application of models is one of the core practices in science, we believe that developing teaching and learning activities on this subject using LS can stimulate both students' and teachers' scientific literacy by helping them to understand theory on model-based reasoning and on biological concept-process models and provide insight in the science behind the lessons they develop.

5.6. Acknowledgements

This paper and the research behind it would not have been possible without the exceptional support of the teachers from the Lesson Study team: Harm Lelieveld, Willemien Hollander and Marieke van Klink. The authors would also like to thank the school Gymnasium Novum and its students for their cooperation in this research.

5.7. Supplementary Figures

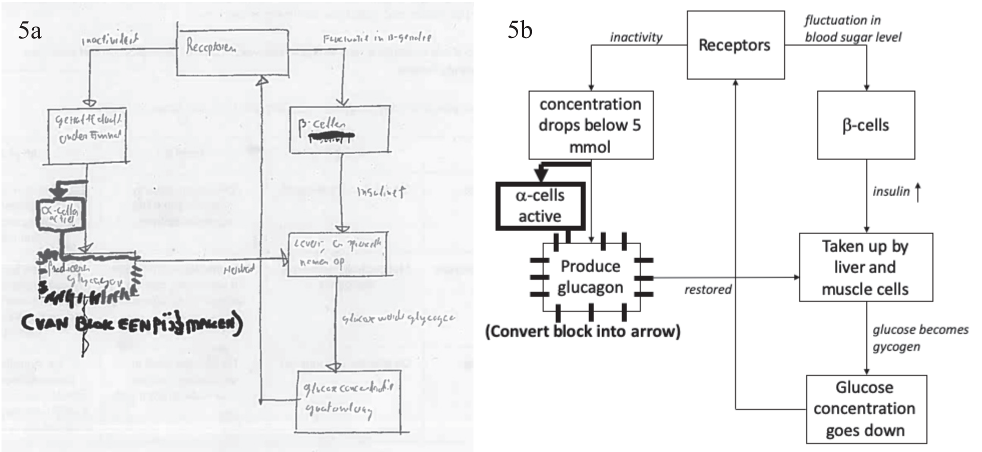


Figure S.1. Model about blood glucose regulation, created by student CS_CC in lesson 2. Figure S.1a shows the original Dutch drawing of the student. Figure S.1b shows a translated version of this model. The model was initially drawn during key activity 2.1 using a pencil and is shown in light grey. The changes that were applied during key activity 2.4 of this lesson have been drawn using a marker and are shown in bold dark grey/black.

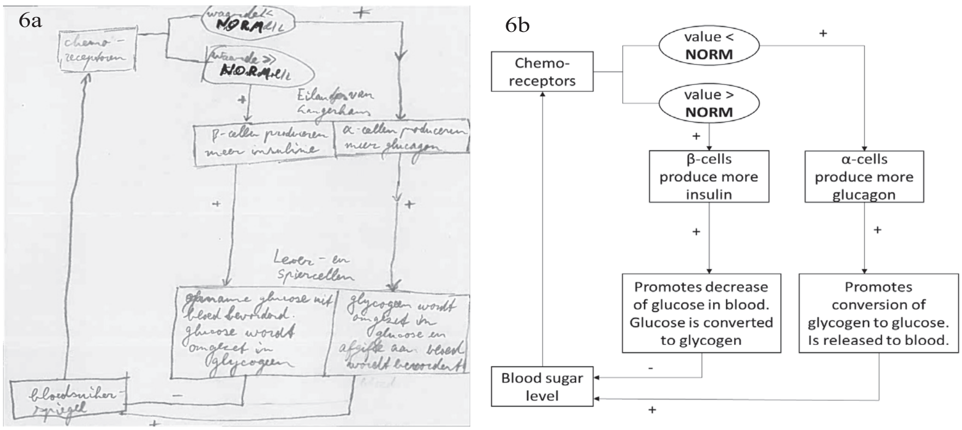


Figure S.2. Model about blood glucose regulation, created by student CM_CA in lesson 2. Figure S.2a shows the original Dutch drawing of the student. Figure S.2b shows a translated version of this model. The model was initially drawn during key activity 2.1 using a pencil and is shown in light grey. The changes that were applied during key activity 2.4 of this lesson have been drawn using a marker and are shown in dark grey/black.

CHAPTER 6

Applying prior meta-modelling knowledge to a VR model of a biological process

This chapter is based on:

Jansen, S., Aung, R., Binte Mohd Taib, S.F., Cai, Y., & van Joolingen, W.R. (accepted for publication). Applying prior meta-modelling knowledge to a VR model of a biological process. In Y. Cai., E. Mangina & S. L. Goei (Eds.), *Mixed Reality for Education*. Springer.

Abstract

The goal of this study was to find out whether students can extend prior meta-modelling knowledge, gained using 2D models, to 3D models. Therefore, we compared a group of students who received prior instruction on models and modelling with a group of students without such preparation and looked for differences in expressed meta-modelling knowledge considering a combination of 2D models and a 3D model. For the 3D model we developed a VR application on the biological process of blood-glucose regulation. Both groups of students worked with the VR application after which they worked on an assignment where they used a combination of 2D models and the VR model to answer questions related to important aspects of meta-modelling knowledge. A pre- and post-test was used to find out whether working with the VR model influenced students' expressed meta-modelling knowledge in general. Results from the assignment showed a higher level of expressed meta-modelling knowledge considering 2D and 3D models for the group of students with prior knowledge than for the group without prior knowledge. The pre- and post-test also showed that working with the VR model led to a higher level of expressed meta-modelling knowledge for the group of students with prior meta-modelling knowledge. For the group of students without this prior knowledge, working with the VR model led to a lower level of expressed meta-modelling knowledge. These results suggest that teaching students about aspects of meta-modelling knowledge and letting them work with these models is an important step in stimulating students' meta-modelling knowledge for both 2D and VR models.

6.1 Introduction

Models and science are inherently intertwined. Models are a central element of scientific inquiry, research and communication, where they are helpful tools for scientists to represent ideas or describe and predict processes that occur in the natural world (Hoskinson, Couch, Zwickl, Hinko, & Caballero, 2014; Svoboda & Passmore, 2013). Creating models is a human enterprise, meant to simplify phenomena and help us make sense of the complex world around us. Models can do this by highlighting certain salient features of a system while minimizing the roles of others (Hoskinson et al., 2014).

Knowledge about the use of models in science is not only useful for scientists, but also for non-scientists. As we have seen for example during the COVID-19 outbreak in 2020, results of research often find their way into society. In the COVID-19 case, models played a very important role in predicting the growth of the pandemic, and communicating to society about this growth to justify the measures taken. Being able to understand such scientific models in daily life, is considered to be a component of scientific literacy (Feinstein, 2011; Grosslight et al., 1991; Oh & Oh, 2011; Schwarz et al., 2009; Sharon & Baram-Tsabari, 2020). Scientific literacy involves the skills that are required for understanding science in everyday life and making personal decisions on socio-scientific issues. An example of such a socio-scientific issue is the vaccination debate in society (Lundström et al., 2012; Roberts & Gott, 2010). Considering scientific models, scientific literacy entails knowledge about models, the creation of models and the use of models (i.e., meta-modelling knowledge) (Grosslight et al., 1991). To stimulate scientific literacy, teaching students about models and the process of modelling as a scientific practice is part of the curriculum in many countries (e.g., the United States Next Generation Science Standards (NGSS Lead States, 2013), the National science curriculum in England GOV.UK, 2015), and the science curriculum in the Netherlands for the subjects physics, chemistry and biology (CvTE, 2018)).

In the science of biology, many different types of models are used. Biological models can range from concrete scale models, such as a model of a human skeleton, to abstract models of complex biological processes, such as the process of photosynthesis. These complex models of biological processes are considered to be abstract, because they contain dynamics such as time and movement, which are often visualized by arrows (Figure 6.1) (Jansen et al., 2019). Since these models consist of several concepts that are connected by these abstract dynamics, they are often called concept-process models (Harrison & Treagust, 2000).

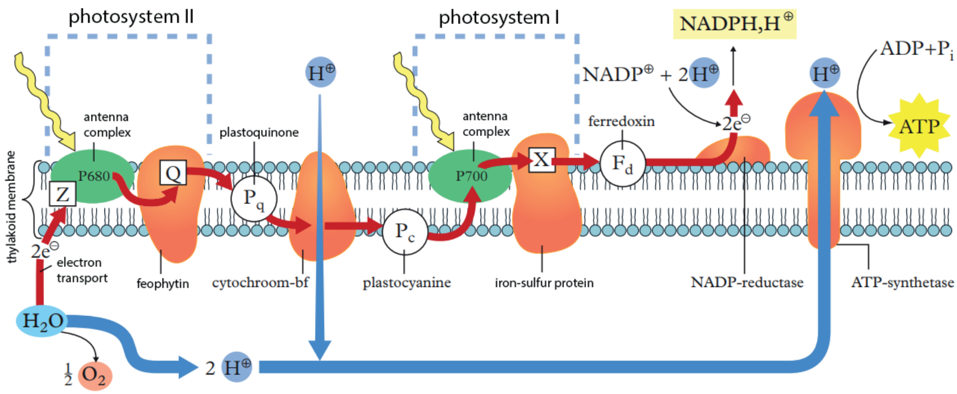


Figure 6.1. A concept-process model of the light reaction of photosynthesis, as used in secondary schools in the Netherlands. The light reaction is the first step in the creation of glucose, the ultimate goal of photosynthesis. Reprinted and translated with permission from Noordhoff Uitgevers, Groningen (Brouwens et al., 2013).

Even though knowledge about models and the process of modelling is part of the curriculum, students are most often not explicitly taught the required meta-modelling knowledge that is necessary to understand models as they are used in science. Most teachers only use models to illustrate biological concepts or phenomena, but neglect teaching the scientific processes of creating and evaluating a model, or using the model to formulate hypotheses (Hoskinson et al., 2014; Windschitl et al., 2008). A solution to this problem can be to incorporate suitable model-based learning approaches into the science curriculum (e.g., Krell et al., 2012; Quillin & Thomas, 2015; Schwarz & White, 2005; Treagust et al., 2002). In practice, model-based learning approaches and model-based inquiries are often reflections and extensions of the scientific method (Windschitl et al., 2008). They typically consist of five steps: (1) observation and data collection, (2) construction of a preliminary model, (3) application, (4) evaluation, and (5) revision of the preliminary model (Fretz et al., 2002).

To evaluate students' meta-modelling knowledge, various frameworks have been developed (e.g., Crawford & Cullin, 2005; Grosslight et al., 1991; Justi & Gilbert, 2003; Louca et al., 2011; Schwarz et al., 2009). In previous research we worked with the framework as described by Upmeier zu Belzen et al., (2019), who combined the frameworks from Crawford & Cullin (2005), Grosslight et al. (1991) and Justi & Gilbert (2003) to create a theoretical framework that can be used as an analytical framework for assessing and investigating students' understanding of models and their use in science. The five main aspects of this framework are *nature of models, multiple models, purpose of models, testing models* and *changing models*. For each of these aspects, several levels of understanding have been formulated, ranging from an initial level of understanding to an expert level of understanding (Level 3). The aspects *nature of models* and *multiple models* reflect on the way models describe phenomena. *Nature of models* focuses on the extent to which a model can be compared to the original. On the lowest level of reasoning, one can think that a model is always an exact copy of the original (Level 1). However, the creator of the model often makes choices, which can aid in emphasizing a specific part of a process. Examples are the colours that are used, the amount of detail that is present and the scale in which certain aspects are depicted (Level 2). At the highest level of reasoning, one can think of a model as a tool to highlight a certain hypothesis about how a process functions (Level 3). The aspect *multiple models* addresses the fact that multiple models can be used to represent the same original. Since no single model can show every detail of a process or explain all existing hypotheses and ideas, models are often simplified to fit the questioners' need and prior knowledge. This results in various different models showing the same process, but with a difference in focus (Level 2) or explaining different ideas (Level 3) (Grünkornet al., 2014; Jansen et al., 2019).

The aspects *purpose of models, testing models* and *changing models* together reflect on the use of models in science: to communicate ideas, to make predictions about future events and to test hypotheses (Grünkorn et al., 2014). The aspect *purpose of models* differentiates in the purpose that models serve in science. Models can be used to show what is known about a process (Level 1), to connect the depicted process to other aspects or processes (Level 2) or to formulate and test hypotheses (Level 3). The aspect *testing models* describes different ways in which models can be validated and the aspect *changing models* shows that models are per definition subject of change, due to for example falsification of a hypothesis about the original. For the aspects *testing models* and *changing models* three levels of understanding have also been described (see Table 6.1).

It was empirically shown by Krell et al. (2012), that the Levels 1, 2 and 3 as mentioned in Table 6.1 reflect an increasing degree of difficulty. Reasoning on Level 1 is not wrong, it is just perceived as a more basic type of reasoning. In general, Level 1 focuses only on the model as a close representation of the original (*model object,* Mahr, 2009). Level 2 has been reached when the model is seen as a medium that has been created based on the original (*model of something,* Mahr, 2009). Level 3 shows an understanding of the use of models in science, in which models are used to test hypotheses and draw conclusions about the original (*model for something,* Mahr, 2009). In contrast to Level 1-3, the initial level that has been formulated for the aspects *testing models, changing* models and *multiple models* cannot be seen as a valid way of reasoning. This level shows a lack of basic understanding of the aspects as mentioned in the framework, rejecting the fact that models can be tested or changed, or that multiple models of the same process exist (Grünkorn et al., 2014).

Since Grünkorn et al. (2014) evaluated this theoretical framework using a variety of biological contexts, this framework was the framework of our choice when working with biological models. However, considering biological concept-process models specifically, we had to extend the framework with categories that are specific for this type of models (Jansen et al., 2019). The resulting framework is shown in Table 6.1, where the categories that are specific for concept-process models are presented in bold. For a detailed description of the categories that are specific for concept-process models, see Jansen et al. (2019).

In previous research we combined the aspects and levels as described in Table 6.1 with both interventions as described by Quillin and Thomas (2015) and pedagogical content knowledge from teachers to develop teaching activities that focused on stimulating meta-modelling knowledge considering biological concept-process models specifically (Jansen et al., 2021a; Jansen et al., 2021b). In our research we mainly focused on the aspects *nature of models* and *multiple models*. These two aspects focus on models as they are presented in for example students' textbooks, without taking the modelling-process that includes testing and changing an existing model into account. Since students most often reason with existing models in class and on tests, it made sense to focus on students' reasoning considering these models. Using the framework we were able to assess the effect of these activities on students' expressed meta-modelling knowledge considering biological concept-process models for the aspects *nature of models* and *multiple models*. Results showed that the activities successfully improved students' expressed meta-modelling knowledge for these two aspects.

In previous research we used existing educational two-dimensional models, such as the one shown in Figure 6.1. However, recently 3D digital models are reaching classrooms in the form of augmented reality (AR) and virtual reality (VR) environments (Velev & Zlateva, 2017). AR is a technology that allows virtual imagery information to be projected onto a live real-world environment (Zhou, Dun, & Billinghurst, 2008). Students can use a phone, tablet or other electronic innovations on which they see the virtual imagery projected onto their environment (Lee, 2012). VR is different from AR in that it delivers an immersive experience, where a three dimensional computer-generated virtual environment is created and the user is able to interact with this environment. In this study we will focus on the use of VR in education.

Shim et al. (2003) argue that VR can be a promising addition to science teaching, since it can stimulate the multi-sensory organs of students, which motivates students and increases their interest in learning activities. Also, it creates the possibility for real-time interaction between both the learner and a computer, and among several different users. Thirdly, the virtual environment enables experimentation to be carried out safely, in an easily controllable environment. This means students can carry out experiments that may otherwise be deemed too dangerous or expensive for the classroom. Finally, as with many other digital environments, students can participate in learning activities at their own comfort and pace. Shim et al. (2003) studied the effect of VR simulations in biology education specifically and found that using VR actively engages students with the biological topic at hand and lets them immerse in the learning environment that is brought to them via the simulation. This experience is shown to be more immersive than the experience that other multimedia options offer.

In this current study we let students work with a VR model of the biological process ‘blood-glucose regulation’, a process that describes the regulation of the concentration of glucose in the blood involving several organs, hormones and enzymes (Ackerman, Gatewood, Rosevear, & Molnar, 1965). Using both two-dimensional concept-process models covering this same process and the VR model, we focus on students’ expressed meta-modelling knowledge for this combination of 2D and 3D model-types. The goal of this study was to find out whether students’ prior meta-modelling knowledge influenced their view on the VR model and whether this knowledge lead to differences in learning result. Therefore, we compared a group who received prior instruction on models and

Table 6.1. Framework to assess students' meta-modelling knowledge of biological concept-process models. The left column shows the five aspects that are important when reasoning with biological models. For each of these aspects up to four levels of understanding have been defined, ranging from an initial level of understanding to an expert level of understanding. Categories in bold have specifically been added to assess students' meta-modelling knowledge of biological concept-process models (Jansen et al., 2019).

	Initial level	Level 1	Level 2	Level 3
Nature of models		- Model as copy - Model with great similarity - Model represents a (non-) subjective conception of the original **- Displays a process, its components and how they are related**	- Parts of the model are a copy - Model as a possible variant - Model as focused representation	- Model as hypothetical representation
Purpose of models		- Model for showing the facts **- Model for showing events**	- Model to identify relationships	- Model to examine abstract/ concrete ideas
Multiple models	- All models are the same - Various models of different originals - Only one final and correct model	- Different model object properties	- Focus on different aspects	- Different assumptions
Testing models	- No testing of models **- Perceiving schoolbooks or their authors as authorities providing absolute truth**	- Testing of material - Testing of basic requirements	- Comparison between original and model - Comparison and matching of original and model	- Testing hypotheses - Testing of hypotheses with research designs
Changing models	- No reason for alterations - Alteration of how different originals are represented	- Alterations to improve the model object - Alterations when there are errors in the model object - Alterations when basic requirements are not met	- Alterations when model does not match the original - Alterations due to new findings about the original	- Alterations due to findings from model experiments **- Alterations when the focus of the model shifts to a different aspect of the process**

modelling with a group without such preparation and looked for differences in students' expressed meta-modelling knowledge after working with the VR model. This comparison was done in order to investigate the following research question:

> *6.1: Does prior instruction in meta-modelling result in differences in applying meta-modelling knowledge to a combination of 2D models and a 3D virtual environment?*

6.2. Method

6.2.1. Participants

In total 88 Dutch eleventh-grade pre-university level students (16 – 18 years old) participated in our study. The group of students who already worked with important aspects of meta-modelling before consisted of 41 students (21 male, 20 female) and is referred to as the prior knowledge-group. These students had received three lessons on model-based reasoning in their previous academic year, in which both the aspects *nature of models, purpose of models* and *multiple models* were discussed, and students gained experience in creating a model of a biological process themselves. The group of students who had not explicitly been introduced to these aspects before consisted of 47 students (20 male, 27 female) and is referred to as the comparison-group. All lessons were taught in the same classroom as where the students' biology lessons usually take place and informed consent was obtained from all students.

6.2.2. Overview of this study

To focus on the effect of prior meta-modelling knowledge on students' expressed meta-modelling knowledge considering a combination of models, both the prior knowledge-group and the comparison-group received the same three lessons. During the first lesson students filled out a pre-test. In the second lesson the students worked with the VR model on blood-glucose regulation, after which they answered questions on paper related to the aspects *nature of models, purpose of models* and *multiple models*. The post-test was filled out during the third lesson.

6.2.3. The VR application on blood-glucose regulation

A VR application was developed, showing the process of blood-glucose regulation. The application can be used on mobile devices and is designed to introduce students to this biological process. This means that prior knowledge about the process of blood-glucose regulation is not necessary to work through the application. Students start in a virtual classroom where someone eats a donut. They then follow the glucose molecules that go from the small intestines, through the cell membrane of the small intestines, into the bloodstream and into a somatic cell. The scenes were chosen based on the different levels of biological organization that are commonly discussed in secondary education when this topic is discussed in class. Figure 6.2 shows stills from the VR application, where different environments are visible.

While working through the application, students answer questions about the process relating to either why something happens the way it does, or they predict what will happen next. All questions are multiple choice, and students receive instant feedback on the answer they choose. Figure 6.3 shows how questions are presented to students in the application.

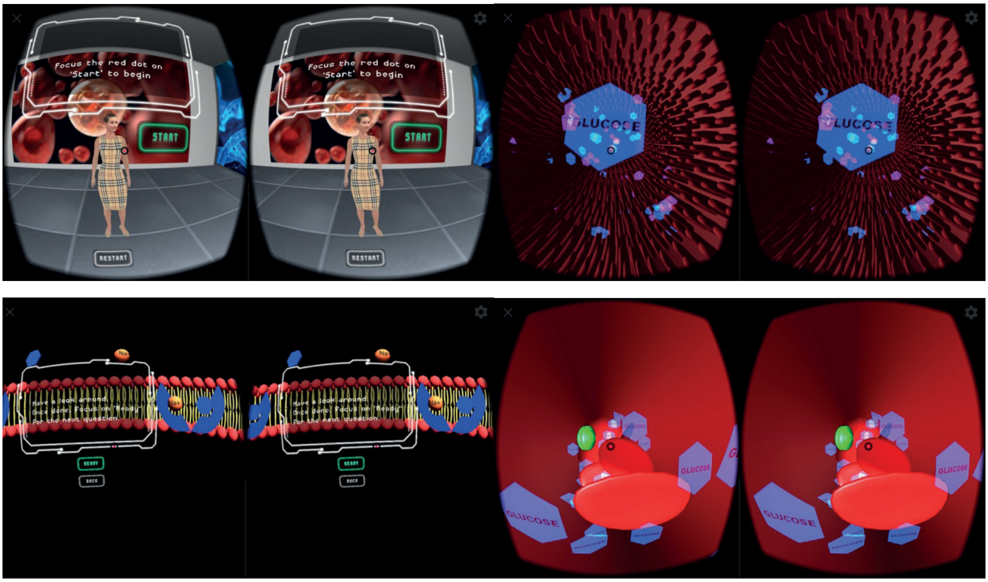


Figure 6.2. Stills from the developed VR application, showing multiple levels of biological organization. The stills are screenshots from a mobile device. When the mobile device is inserted into VR-goggles, the two circles merge and the user experiences a 3D environment.

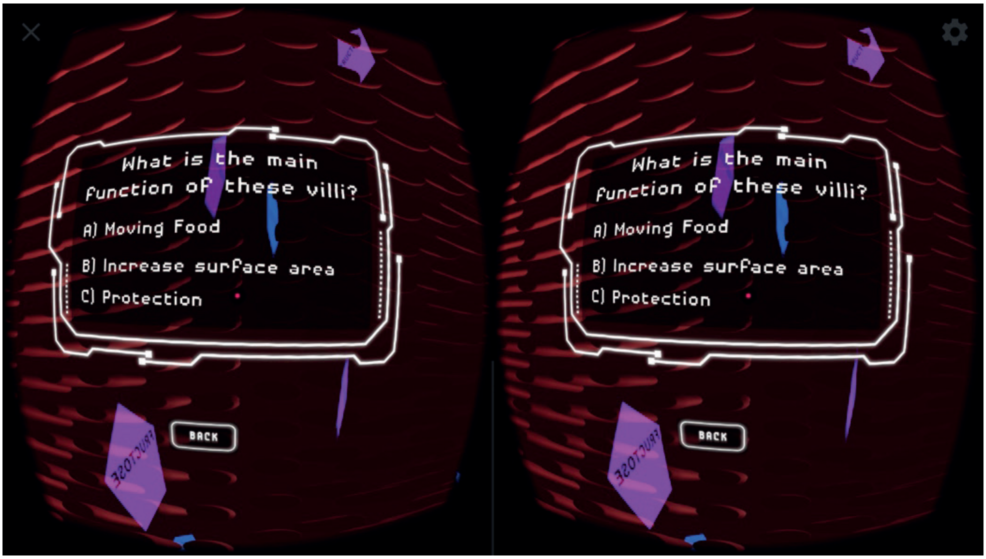


Figure 6.3. A multiple-choice question as presented to students in the VR application. The red dot in the centre of each of the circles can be used to choose an answer.

After selecting an answer using the 'red dot' in the centre of their vision, students hear whether the answer is correct or not. They also hear an explanation about why the answer is correct or not. When students do not answer correctly, they choose again. This cycle repeats itself until a student chooses the correct answer. The following text shows the way the example in Figure 6.3 is introduced to the students and what kind of feedback they receive for each answer they choose.

> *Teacher: the lumen of the small intestines is lined with villi. These are the tentacle-like structures that you see all around you. What is the main function of these villi?*
>
> *A: They help digestion by moving/waving the food in the right direction (text on screen: moving food)*
>
> *B: They increase the surface area for the uptake of nutrients (text on screen: increase surface area)*
>
> *C: They protect the intestines from potentially damaging particles from your food (text on screen: protection)*

In coherence with the given answer, the students hear a response from the teacher in the application:

A: Sorry, that is incorrect. Muscle contractions of the small intestines move the food in the right direction. Let's try again!

B: You're right, awesome! A larger surface area increases the chance of nutrients being taken up by your body.

C: Sorry, that is incorrect. The shape of the villi does not contribute to this form of protection. Let's try again!

To find out whether there is a difference between the prior knowledge-group and the comparison-group considering the way students work through the application, the first answer that students choose is registered for each question and for each student separately.

6.2.3.1. Pilot

The VR application was tested in a pilot session with 10 pre-university education students with a major in biology (16 - 17 years old) (Figure 6.4).

Figure 6.4. Testing the VR application. Students use their own mobile device, which they inserted into a Google cardboard device.

Students worked through the application, after which they filled out an open-ended questionnaire where they could comment on the quality of the graphics, the quality of the audio, the given interaction possibilities, the way the blood-glucose regulation was explained in the application, and whether the questions in the application fitted the knowledge level of the students. Students were also given the opportunity to write down any other remarks or recommendations they had considering the use of the application. Results from the pilot study indicated that students very much enjoyed working with the application, that the audio worked well, that the feedback helped them understand the process better, that the process was easy to follow, and that the questions fitted the knowledge level of the students well. Two things did not work well according to the students: the graphics were blurry or slow on older phones, and the red dot that was implemented to interact with the virtual environment was difficult to see when the environment was red in colour as well (inside the bloodstream). To make the red dot visible at all times, we placed a black circle around the red dot. Considering the blurriness, we decided not to compromise on the graphics. Instead, we decided to bring a set of mobile phones on which we knew the application worked well to the classes in which we carried out our research.

6.2.4. Developed student assignment

To find out how students compared working with 2D models and the VR model, students received an assignment on paper after working with the application in which they were asked to use both types of models when answering questions related to the aspects *nature of models, purpose of models* and *multiple models*. Students were presented with four different models on blood-glucose regulation varying in level of biological organization and abstractness. Three of these models were 2D and one was the developed three-dimensional VR model. These four models were printed on paper for students to use while answering the questions during the assignment. For the VR model, a still from the application was printed and students were reminded that for this particular model they had to think of the complete VR simulation when answering the questions and not just look at the printed still.

The assignment consisted of seven tasks, which were based on the activities we developed during previous research that were used to introduce students to the aspects *nature of models, purpose of models* and *multiple models*. For a description of these activities and the theoretical background behind the tasks, see Jansen et al. (2021a). The seven tasks were as follows:

1. Students had to name differences between the four models (relating to the aspect multiple models)
2. Students had to categorize the differences that they wrote down for question 1 (relating to the aspect multiple models)
3. Students had to match given aims with the four different models and explain their choices (relating to the aspect purpose of models)
4. Students had to explain whether all aims from question 3 could be met using only the VR model (relating to the aspect purpose of models)
5. Students had to explain whether all aims from question 3 could be met using only one of the 2D models (relating to the aspect purpose of models)
6. Students had to name choices that the creator of the models made in order to meet the prospected aim of these models (relating to the aspect nature of models)
7. Students had to explain whether they thought that eventually one ultimate model for the description of blood-glucose regulation would suffice (relating to the aspect nature of models)

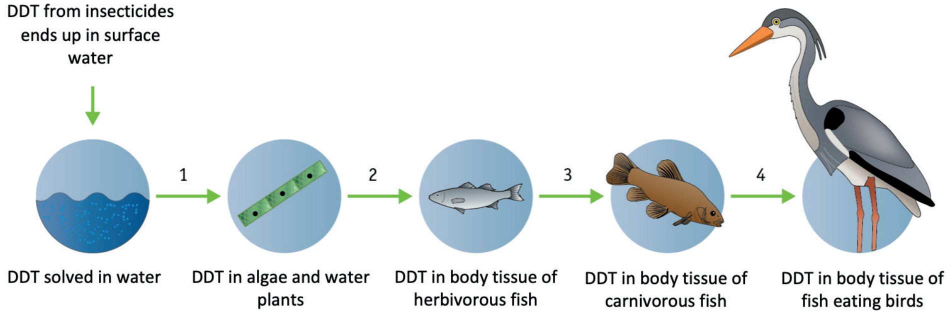
6.2.5. Pre- and post-test

To measure the influence of having prior knowledge about the aspects *nature of models*, *purpose of models* and *multiple models* on students' expressed meta-modelling knowledge for these three aspects we used an online test. This test was previously developed based on work from Krell (2019) and assesses students' level of expressed meta-modelling knowledge for biological concept-process models specifically (van Montfort, 2019; Jansen et al.,(2021c). The test contains statements, relating to six different biological concept-process models. For each of these models a statement on Level 1, Level 2 and Level 3 is formulated for the aspects *nature of models* and *multiple models* (see Table 6.1 for a description of the levels). The test only focuses on two aspects, since including all aspects would generate too many questions for students to answer. The choice for the aspects *nature of models* and *multiple models* was made since these aspects reflect on the students' meta-modelling knowledge considering existing concept-process models, resembling both the way students are presented with models during tests in school and the way students encounter scientific information in daily life. In the test, students have to mention whether they agree or disagree with the given statements. Students' agreement with the statements is interpreted as their expressed meta-modelling knowledge for the aspects *nature of models* and *multiple models*

(Krell, 2019; Jansen et al., 2021a). Figure 6.5 shows a translated example of one of the closed-question tasks in this test, containing a model showing the process of bioaccumulation. In this case the task reflects on the aspect *multiple models*. A link to the online test can be found in Appendix C.

Example item

Frank is doing research on water contamination. He discovered that in areas where the water is contaminated with DDT (a chemical insecticide), many fish eating birds are dying. He creates the following model.



Maud is also doing research on water contamination with DDT. She creates a model of the same process. When can the model that Maud creates be defined as a different model than Franks' model?

Answer for each of the statements if you agree (yes) or disagree (no)

When Maud creates a model for an area that is contaminated with DDT, but where different species of fish are living.
[ ] Yes
[ ] No

When Maud creates a model where the focus is more on step 2 of the process.
[ ] Yes
[ ] No

When Maud has different thoughts on how DDT ends up in fish eating birds and creates a model reflecting these thoughts.
[ ] Yes
[ ] No

Figure 6.5. An example of one of the tasks in the test, showing a model of bioaccumulation. The aspect of focus in this example is *multiple models* and each statement represents a level of understanding in the order: Level 1, Level 2 and Level 3 (the order of the levels differed between different tasks). The model present in this task is reprinted and translated with permission from Bos et al. (2012).

6.2.6. Student interviews

To find out how students experienced the intervention and to substantiate the results from the student material and the pre- and post-test, four students from the comparison-group and six students from the prior knowledge-group were interviewed after the completion of the lesson series (three lessons). Interviews were recorded and lasted approximately 20 minutes. The students were asked to reflect on both their experience with the VR application and the student material that contained questions about the aspects *nature of models, purpose of models* and *multiple models*. The interview questions were related to whether students had any experience working with VR before the intervention, whether they would recommend using this VR application in class, in what way 2D models from their biology textbook are usually discussed in class, whether their biology teacher discusses the use of models in biology in general (and if so, in what way), which question from the student material they thought was difficult and which question they thought was easy, and whether they thought it would be good to spend more time in class on the use of models in biology in general. The prior knowledge-group was also asked whether they had used the theory about important aspects of meta-modelling knowledge that they learned about in their previous year while answering the questions in the student material, and whether they thought having knowledge about these aspects helped them understand the models that are used in biology education better.

6.2.7. Data management

6.2.7.1. VR application

We used the student answers from the VR application on the topic blood-glucose regulation to determine whether there was a difference in prior knowledge considering the process of blood-glucose regulation between the two groups on this topic. The percentage of students that answered the questions in the VR application correctly on their first try was determined for each question separately. Using an unpaired t-test we determined whether differences between the two groups could be considered significant.

6.2.7.2. Student assignment

The student assignment contained questions relating to the aspects *nature of models, multiple models* and *purpose of models*. Student answers were coded using the categories, aspects and levels as described in Table 6.1. Ten percent of the student work was coded by a second independent coder, resulting in 84% agreement between the first and second coder. After discussing differences, 100% agreement was reached. The average agreement score (AAS) per aspect and Level was determined for both the prior knowledge-group and the comparison-group. Using an unpaired t-test we calculated whether the differences in student answers could be considered significant.

6.2.7.3. Pre- and post-test

The results from the pre- and post-test were analysed to determine the influence of having prior knowledge about the aspects *nature of models, purpose of models* and *multiple models* on students' expressed meta-modelling knowledge for the aspects *nature of models* and *multiple models.* The results of the test showed for six different biological concept-process models whether students agreed with statements on Level 1, Level 2 and Level 3 for the aspects *nature of models* and *multiple models.* The agreement score for all levels and aspects was determined separately for each student. Since the pre- and post-test contained six tasks per aspect, and each task contained three statements (on Level 1, Level 2 and Level 3), the agreement scores per aspect ranges from 0 to 6, where 0 means disagreement with all statements, and 6 means agreement with all statements. After determining the agreement score per student per question, the AAS per aspect (*nature of models* and *multiple models*) for students in the prior knowledge-group and the comparison-group was determined. Using a paired t-test we calculated whether the difference between the number of times a student agreed with statements on a certain level in the pre-test and post-test could be considered significant. Using an unpaired t-test we calculated whether differences between the prior knowledge-group and the comparison-group could be considered significant.

6.2.7.4. Student interviews

Student interviews were transcribed verbatim and scanned for utterances related to students' experience with the VR application, their view on the use of VR in class and their view on the use of models in education. These utterances were only used to substantiate the results from the pre- and post-test and from the student assignment.

6.3. Results

6.3.1. VR application

The VR application on blood-glucose regulation contained 11 questions. The percentage of students that answered each question correctly on their first try was determined for both the experimental group and the control group. Using an unpaired t-test we determined for each question whether there was a difference in the number of correct answers between the experimental group and the control group. Table 6.2 shows the results per question, where no significant differences between the prior knowledge-group and the comparison-group were found. As shown in Table 6.2, the number of registered student answers drops for each question in both the prior-knowledge group and the comparison-group. Even though almost all students finished working with the app, not all student answers were registered. This is because the app only registers a student answer the first time the app is started by that student. When a student exits the app and restarts the app, no new student answers are registered. Students accidentally pressing the exit button on their phone, WiFi problems causing the app to crash, or students switching from phones after a few questions because they felt their phone was 'too slow' caused the drop in registered number of student answers.

Table 6.2. Results per question for the VR application. Only those answers that were registered were taken into account (*n* = number of registered student answers, p-group = prior knowledge-group, c-group = comparison-group).

Question	P-group (n)	C-group (n)	Correct answers p-group (%)	Correct answers c-group (%)	t	p
1	47	64	46.8	62.5	1.64	.10
2	40	60	82.5	78.3	.51	.61
3	30	59	76.7	72.9	.39	.70
4	22	55	86.4	92.7	.77	.45
5	17	54	58.8	59.3	.03	.98
6	15	51	93.3	96.1	.38	.71
7	12	49	25.0	51.0	1.74	.09
8	13	41	92.3	80.5	1.19	.24
9	11	35	100.0	100.0	n/a	n/a
10	10	33	80.0	97.0	1.24	.25
11	10	28	60.0	60.7	.03	.97

6.3.2. Student assignment

Student answers on the assignment were scored using the framework as shown in Table 6.1. Answers to questions relating to the same aspect of meta-modelling knowledge (*nature of models, multiple models* and *purpose of models*) were combined to get an overview of students' reasoning per aspect. An unpaired t-test was carried out to compare results from the prior knowledge-group and the comparison-group. Table 6.3 shows these results, showing that the prior-knowledge group more often scores a Level 2 answer for the aspects *nature of models* and *multiple models* than the comparison-group. The comparison-group more often scores a Level 1 answer for the aspects *nature of models* and *multiple models*.

Table 6.3. Results for the student assignment per aspect and level (*n* = number of scored student answers, *M* = mean (ranging from 0 – 1), p-group = prior knowledge-group, c-group = comparison-group).

		P-group (n)	C-group (n)	M p-group	M c-group	t	p
Nature of models	Level 1	90	109	.71	.86	2.59	.01
	Level 2	90	109	.61	.32	4.24	<.01
	Level 3	90	109	.03	.01	1.14	.25
Multiple models	Level 1	136	193	.90	.97	2.70	.01
	Level 2	136	193	.53	.33	3.71	<.01
	Level 3	136	193	.01	.00	1.00	.32
Purpose of models	Level 1	134	188	.93	.96	1.18	.24
	Level 2	134	188	.07	.04	1.40	.16
	Level 3	134	188	.00	.01	1.00	.32

6.3.3. Pre- and post-test

Students' agreement with statements on the pre- and post-test can be directly related to their expressed meta-modelling knowledge for the aspects *nature of models* and *multiple models*. Students' average agreement score (AAS) for each aspect and level was determined. Since students were presented with six statements for each level and aspect, students' AAS ranges from 0 – 6. Results for the pre- and post-test were compared for both the prior knowledge-group and the comparison-group using a paired t-test. Table 6.4 shows these results. For the prior knowledge-group we found no significant differences in AAS between the pre-test and the post-test. In the control group the difference in AAS between the pre-test and the post-test for statements relating to *multiple models* Level 1 and *multiple models* Level 3 could be considered significant. Students' AAS was higher in the post-test than in the pre-test for statements relating to *multiple models* Level 1, and lower in the post-test than in the pre-test for statements relating to *multiple models* Level 3.

An unpaired t-test was used to find out whether there are differences in AAS between the prior knowledge-group and the comparison-group for both the pre-test and the post-test. Table 6.5 shows the results for this comparison. The difference in AAS between the prior knowledge-group and the comparison-group for the aspect *nature of models* could be considered significant for statements relating to Level 3 in the pre-test. The difference in AAS between the prior knowledge-group and the comparison-group for the aspect *multiple models* could be considered significant for statements relating to Level 2 and Level 3 in the post-test. For all these differences the AAS for the prior knowledge-group was higher than the AAS for the comparison-group.

Table 6.4. Comparison between the pre-test and the post-test for both the prior knowledge-group and the comparison-group (AAS = average agreement score).

		Prior knowledge-group				**Comparison-group**			
		AAS pre-test	**AAS post-test**	***t***	***p***	**AAS pre-test**	**AAS post-test**	***t***	***p***
Nature of models	Level 1	4.31	4.42	.42	.68	4.27	4.55	.96	.34
	Level 2	4.89	5.18	1.11	.27	4.64	4.47	.57	.57
	Level 3	5.13	5.16	.10	.92	4.39	4.59	.60	.55
Multiple models	Level 1	1.47	1.84	1.31	.19	1.10	2.07	4.12	<.01
	Level 2	4.18	4.87	1.47	.15	3.59	3.75	.39	.70
	Level 3	5.63	5.68	.27	.79	5.41	4.80	2.62	.01

Table 6.5. Comparison between the prior knowledge-group and the comparison-group for both the pre-test and the post-test (AAS = average agreement score, p-group = prior knowledge-group, c-group = comparison-group).

		Pre-test				Post-test			
		AAS p-group	**AAS c-group**	***t***	***p***	**AAS p-group**	**AAS c-group**	***t***	***p***
Nature of models	Level 1	4.31	4.27	.19	.85	4.42	4.55	.41	.68
	Level 2	4.89	4.64	1.01	.31	5.18	4.47	2.50	.01
	Level 3	5.13	4.39	2.75	<.01	5.16	4.59	1.68	.10
Multiple models	Level 1	1.47	1.10	1.64	.11	1.84	2.07	.80	.43
	Level 2	4.18	3.59	1.28	.21	4.87	3.75	2.76	<.01
	Level 3	5.63	5.41	1.15	.25	5.68	4.80	3.78	<.01

6.4. Discussion

The aim of this study was to find out what the influence of prior experience with meta-modelling aspects was on students' expressed meta-modelling knowledge considering a combination of biological 2D models and the VR model on blood-glucose regulation. We compared a group of students who previously received three lessons on aspects of meta-modelling knowledge (prior knowledge-group) with a group of students who did not receive these lessons (comparison-group). Both groups worked with the same VR model on blood-glucose regulation, after which they answered questions relating to both 2D models and the VR model focusing on different aspects of meta-modelling knowledge.

Multiple questions related to the biological process blood-glucose regulation were incorporated in the developed VR application. These questions were not related to aspects of meta-modelling knowledge and were meant to see whether students' knowledge on the biological topic was comparable between the experimental group and the control group. No significant differences were found between student answers from both groups. Despite the low number of registered student answers for the second half of the questions in the VR application, we assumed that no difference in prior knowledge considering the topic of blood glucose regulation between the prior knowledge-group and the comparison-group existed.

Both groups filled out an online pre- and post-test to find out students' expressed level of meta-modelling knowledge relating to the aspects *nature of models* and *multiple models*. In this test students have to agree or disagree with statements related to the three levels of meta-modelling knowledge (Table 6.1) for both of these aspects. All three levels show valid ways of reasoning, but they increase in difficulty. Ideally students should be able to reason on all three levels as described in Table 6.1 (Krell, 2019). When comparing the results from the prior knowledge-group and the comparison-group, we see that in the pre-test the prior knowledge-group significantly more often agreed with statements related to the highest level of reasoning (Level 3) for the aspect *nature of models* than the comparison-group. In the post-test the prior knowledge-group more often agreed with statements relating to Level 2 and Level 3 for the aspect *multiple models* than the comparison-group. The prior knowledge-group's higher agreement score relating to the higher levels of reasoning suggests that the lessons they received in their previous academic year influenced their expressed meta-modelling knowledge.

Looking at the difference in results between the pre- and post-test for both the prior knowledge-group and the comparison-group, we see that while there is no significant difference in student results for the prior knowledge-group, the comparison-group significantly more often agreed with statements on Level 1 for the aspect *multiple models* and significantly less often with statements on Level 3 for the aspect *multiple models* in the post-test than in the pre-test. The lower AAS for Level 3 statements in the post-test for the aspect *multiple models* shows that just answering questions that are related to the aspects of meta-modelling knowledge without receiving any theory on these aspects does not trigger a Level 3 type of reasoning. This suggests that in order to stimulate this level of reasoning, students need practice related to this aspect and level of meta-modelling knowledge. The increase in AAS for *multiple models* Level 1 statements could be due to the assignment where students were shown different models related to the same biological topic (blood-glucose regulation) and had to name differences between these models. Since a Level 1 type of reasoning for this aspect focuses on aesthetic differences, such as differences in colour, having seen these differences during the lesson might have been enough to trigger this level of reasoning. The higher agreement score from the prior knowledge-group for Level 2 statements in the post-test suggests that working with the different models and/or being confronted again with questions relating to aspects of meta-modelling knowledge might have freshened up a Level 2 type of reasoning in students for this aspect. The following student quote shows that students from the prior knowledge-group were not reminded of the theory on model-based reasoning outside the three lessons they received in their previous year. This means that the current lesson, where they had to work with three of the aspects (*nature of models, purpose of models* and *multiple models*), was the first time since these three lessons in their previous year that they were reminded of these aspects.

> *E2: Usually the teacher shows a picture of a model on the screen and then we just go through the model together. So, the teacher says something like, you have this substance and that goes in here, then this happens, and then that. So just a bit of explanation about the process, but nothing more. Apart from those three lessons last year, nothing is said about models in general.*

The assignment on paper that students received after working with the VR application consisted of open-ended questions related to the aspects nature of models, purpose of models and multiple models. Students used both 2D models and the VR model to answer these questions. Results show that the prior knowledge-group significantly more often answered with a Level 2 type of reasoning for the aspects nature of models and multiple models, and the comparison-group significantly more often answered with a Level 1 type of reasoning for these aspects. Since the questions were open-ended, students most often only wrote down one type of answer. This means that the answer they gave can be interpreted as their preferred level of reasoning, and that the students in the prior knowledge-group preferred a higher level of reasoning for the aspects nature of models and multiple models than the comparison-group. The first question, where students had to write down differences between the four models of blood-glucose regulation, can be seen as an exception. Since students had to write down at least six differences, there was room for answers on multiple levels of reasoning for the aspect multiple models. Results show that students from the prior knowledge-group often do not mention a Level 1 type of reasoning (aesthetic difference) among these differences. This might be caused by the lessons they received in their previous academic year on the different levels of meta-modelling knowledge. Having knowledge about the higher levels of reasoning might cause students to reject the lower levels of reasoning, even though these levels are still valid ways of reasoning. The following student quote from the prior knowledge-group illustrates this way of thinking:

> *E1: I thought the first question was very difficult, because you had to name six differences between the models. And then I think, well, a difference in colour is a difference, but is that difference noteworthy? That made it difficult for me, because now I looked for differences that I thought would really be noteworthy, but I couldn't find six of those.*

Both the prior knowledge-group and the comparison-group mention the VR model and the 2D models when answering the questions on the student assignment. This indicates that the lessons about important aspects of meta-modelling knowledge that the prior knowledge-group received, which only contained 2D models, did not lead to a preference in type of models when answering questions that are related to these aspects. Our results also indicate that students' knowledge about important aspects of meta-modelling knowledge related to 2D models also leads to higher levels of expressed meta-modelling knowledge when working with a combination of 2D models and the VR model. However, since this study used a combination of 2D models and the VR model, future research should focus on the way students apply their meta-modelling knowledge to VR models only.

Student interviews revealed how students experienced using VR in class. Consistent with literature on this topic (Shim et al., 2003), students mentioned that they enjoyed working with the VR application. As shown by the following student quote, students found working with the app engaging.

> *C1: I enjoyed the lesson, it was more fun than a traditional lesson. There was an interactive aspect in the application. I like that a lot, that's the way I like to learn something. And you're constantly busy, you cannot really get distracted. It was also relatively quiet in class, that's usually not the case. So, it really had a positive impact on the way of learning and on the ambiance in class. It motivated me more to learn, because when I just read something it feels like I'm quickly filled with information and then I get tired and bored. This interactive aspect works much better for me.*

Both the comparison-group and the prior knowledge-group respond in a positive way when asked whether they would recommend using VR in biology education. They mention that using VR would bring more variety in the lessons, that the virtual environment keeps them focused on the biological topic, and that they appreciate the interactive element in the VR application because it results in a more personal experience where each student can go through the application at their own pace. In line with the research by Shim et al. (2003), one of the students mentioned an advantage of using VR for biology lessons specifically:

> *E3: Especially for biology lessons I can see the advantages. Because you can really look inside an organ. Not in the way we do in a practical lesson with a knife, but that you can also see what happens inside an organ. I think that is pretty useful. Of course, there are other ways in which you can sort of learn about that, but those are usually images, or you have to use a knife to look inside an organ, or it's a static model.*

Even though our study showed that students are able to work with both 2D models and the VR model, it has to be noted that the VR model differs from the 2D models in more than one way. Not only provides the VR model a 3D instead of a 2D experience, it also brings the opportunity to merge multiple levels of biological organization into a single application and it creates an interactive environment where dynamics such as time and movement can be visualized without the use of arrows. This raises the question whether the VR application should be interpreted as a single model, or as multiple models in a single application. As mentioned by Ainsworth (2006), the use of multiple models can be beneficial for learning when a single representation would be too complicated if it presented all the information or if the information is on radically different scales. This justifies the use of multiple models, or sub-models, in the VR application and shows that multiple 2D models are necessary for students to understand a biological process on multiple levels of biological organization. When students worked with the VR model, they actually worked with multiple sub-models of the same biological process. In the VR model students shift between different levels of biological organization, meaning that they see the process on different scales. In similar situations where students work with 2D models, including in the current study, they usually work with multiple 2D models showing the same levels of biological organization as we used in the VR model. Therefore, we believe the multi-model nature of the VR model did not have an impact on our study. The following student quote shows how a student from the prior knowledge-group looked at this aspect:

> *E4: I really enjoyed the application, I felt it was a good way to use multiple sub-models, so to say. As if you can zoom in from one model into another aspect [of the process] and see more details. That made things very clear to me.*

This study showed that students who previously learned about important meta-modelling aspects using 2D models are able to extend this knowledge to a VR model. Students' prior knowledge considering meta-modelling aspects also contributes to a higher level of expressed meta-modelling knowledge in the post-test. These results suggest that teaching students about aspects of meta-modelling knowledge and letting them work with these models is an important step in stimulating students' meta-modelling knowledge for both 2D and VR models.

CHAPTER 7

General discussion

7.1 Introduction

In this dissertation we focused on students' meta-modelling knowledge considering a specific type of models that is often used in biology education: concept-process models. Concept-process models show different types of interactions and dynamics and can not only be used to describe and simplify the real-world situation, but also to formulate hypotheses and predict future events (Harrison & Treagust, 2000). Consequently, these models can be seen as a suitable option for teachers to use when stimulating students' meta-modelling knowledge. The aim of this research project was to investigate secondary school students' meta-modelling knowledge considering biological concept-process models and develop teaching and learning activities to stimulate model-based reasoning with this specific type of biological models. The guiding research question for this dissertation was:

> *How can secondary school students' meta-modelling knowledge considering biological concept-process models be fostered?*

We designed a test to assess students' meta-modelling knowledge considering biological concept-process models. This test was first used to investigate the current state of students' meta-modelling knowledge in the Netherlands. Results from this study were used to design teaching and learning activities to foster students' meta-modelling knowledge, where the developed test functioned as a pre- and post-test to find out what the effect of the developed teaching and learning activities was on students' expressed meta-modelling knowledge. The teaching and learning activities were designed and studied using Lesson Study as a research approach. This resulted in three lessons containing multiple key activities to foster students meta-modelling knowledge considering two-dimensional (2D) biological concept-process models. The three lessons were taught in the course of one schoolyear. In the following schoolyear, a virtual reality (VR) concept-process model was developed to find out whether students were able to apply their gained meta-modelling knowledge on a combination of 2D concept-process models and the VR model.

7.2 Research overview and main findings

In **Chapter 2** we discussed the applicability of a framework containing important aspects of meta-modelling knowledge as described by Grünkornet al. (2014) to biological concept-process models. Students from four different schools in the Netherlands were interviewed to evaluate the applicability of this framework. This study was guided by the following research question:

> *2.1: To what extent can the described framework be used to assess students' understanding of biological concept-process models*?

Most student answers fitted the descriptions as presented in the framework. This was in line with our expectations, since the framework had previously been evaluated using various biological contexts (Grünkorn et al., 2014). However, four additional descriptions were necessary to describe students' reasoning with biological concept-process models. These additions resulted in the framework as shown in Table 7.1, where the added descriptions are displayed in bold. The framework shows five important aspects of meta-modelling knowledge (*nature of models, multiple models, purpose of models, testing models* and *changing models*) and four different levels of understanding, ranging from an initial level of understanding to an expert level of understanding (Level 3). Especially the added descriptions for the aspects *nature of models* and *purpose of models* reflect on the dynamics and interactions that are present in concept-process models. The addition for the aspect *nature of models* describes the visualized interactions in concept-process models, and the addition for the aspect *purpose of models* reflects on the idea that processes are dynamic events or phenomena instead of static concepts.

Student answers in this study also presented us with the possibility to give an indication of the current level of students' expressed understanding of biological concept-process models. We found that in general students reached Level 2 for most aspects within this framework. However, an interesting distinction in students' expressed understanding could be made. Students reached the highest levels within this framework more often when asked about a model on the molecular/cellular level ((sub)microscopic level), than when asked about a model on the level of ecological community (macroscopic level), such as the Oostvaardersplassen (nature reserve in the Netherlands).

Table 7.1. Framework to assess students' meta-modelling knowledge of biological concept-process models. The left column shows the five aspects that are important when reasoning with biological models. For each of these aspects up to four levels of understanding have been defined, ranging from an initial level of understanding to an expert level of understanding. Categories in bold have specifically been added to assess students' meta-modelling knowledge of biological concept-process models.

	Initial level	Level 1	Level 2	Level 3
Nature of models		- Model as copy - Model with great similarity - Model represents a (non-) subjective conception of the original **- Displays a process, its components and how they are related**	- Parts of the model are a copy - Model as a possible variant - Model as focused representation	- Model as hypothetical representation
Purpose of models		- Model for showing the facts **- Model for showing events**	- Model to identify relationships	- Model to examine abstract/ concrete ideas
Multiple models	- All models are the same - Various models of different originals - Only one final and correct model	- Different model object properties	- Focus on different aspects	- Different assumptions
Testing models	- No testing of models **- Perceiving schoolbooks or their authors as authorities providing absolute truth**	- Testing of material - Testing of basic requirements	- Comparison between original and model - Comparison and matching of original and model	- Testing hypotheses - Testing of hypotheses with research designs
Changing models	- No reason for alterations - Alteration of how different originals are represented	- Alterations to improve the model object - Alterations when there are errors in the model object - Alterations when basic requirements are not met	- Alterations when model does not match the original - Alterations due to new findings about the original	- Alterations due to findings from model experiments **- Alterations when the focus of the model shifts to a different aspect of the process**

To obtain a reliable picture of the current state of Dutch secondary students' expressed meta-modelling knowledge considering biological concept-process models we needed a test that could easily be distributed among many students throughout the Netherlands. The design of this test and its application are described in **Chapter 3.** In this design we took the results from Chapter 2 considering the added descriptions and the distinction in students' expressed level of understanding between models on the (sub)microscopic level and macroscopic level into account. Based on work by Krell (2019), we also took the presence or absence of a purpose of a model in the context in which models in the test are presented to students into consideration. The following research questions were addressed in this study:

> *3.1: How does students' expressed meta-modelling knowledge, related to the aspects nature of models and multiple models, depend on the presence of different kinds of explicit modelling-purposes?*
>
> *3.2: How does students' expressed meta-modelling knowledge, related to the aspects nature of models and multiple models, differ between contexts showing macroscopic and (sub)microscopic concept-process models?*

The developed test focused on the first two aspects within the framework: *nature of models* and *multiple models*. Results showed that providing students with an explicit modelling purpose had no significant effect on their expressed level of meta-modelling knowledge considering biological concept-process models (RQ 3.1). Results for RQ 3.2 were in line with our observations from Chapter 2 considering the difference in students' reasoning between models on the (sub) microscopic level and the macroscopic level, showing that students more often agreed with higher level statements (Level 2 and Level 3) when presented with (sub)microscopic models than when presented with macroscopic models.

The results from the study described in Chapter 3 indicated focal points for the development of model-based learning approaches to stimulate students' meta-modelling knowledge considering biological concept-process models. Ideally, students should be able to reason on Level 1, Level 2 and Level 3 for all aspects as described in Table 7.1. Even though a Level 1 type of reasoning is considered to be less complex than a Level 3 type of reasoning, all three levels display a valid way of reasoning (Krell & Krüger, 2016). In an ideal situation, students are aware of the different levels and are able to shift between these levels according to the question at hand. This means that, according to the results as described in Chapter 3, extra attention should be given to Level 3 for the aspect *nature of models*, and to Level 1 and Level 3 for the aspect *multiple models*.

Chapter 4 discusses the development of the first lesson in a series of three lessons to support students' meta-modelling knowledge considering biological concept-process models. Lesson Study (LS) was used as a research approach for the development of this first lesson. Since we adapted the LS approach by adding researchers to a team of teachers that designed the lesson, the development of this lesson served two goals: to introduce students to important aspects of meta-modelling knowledge (Table 7.1) and to investigate whether the adapted version of LS had potential to be used as a research approach by the educational research community. The following two research questions were addressed in this study:

> *4.1: To what extend does the developed lesson successfully familiarize students with important levels and aspects that are associated with model-based reasoning?*
>
> *4.2: How do teachers and researchers experience using LS as a research approach?*

The developed lesson successfully familiarized students with the three important aspects of meta-modelling knowledge that were introduced to the students (*nature of models, multiple models* and *purpose of models*). In this lesson students were confronted with four models of the same biological process. They named differences between these models (Key activity 1, relating to the aspect *multiple models*), matched aims of a model to the different models (Key activity 2, relating to the aspect *purpose of* models), and formulated choices that the creator of the model had made in order to meet the aim of the model (Key activity 3, relating to the aspect *nature of* models) (RQ 4.1). Considering the use of LS as a research approach, three main focal points could be formulated (RQ 4.2): 1) make sure that the teachers support the research question that the researchers bring into the LS-cycle, 2) take into account that the lesson is supposed to answer a research question which might cause extra stress for the teachers in a LS-team and 3) state the role of both researchers and teachers in a LS-team clearly at the beginning of the design process.

The focal points from Chapter 4 were considered when Lesson Study was used again as a research approach to develop two additional lessons to stimulate students' reasoning with biological concept-process models. **Chapter 5** describes the design process and execution of these two lessons, in which the first lesson addresses the skill to read and write visual or symbolic language (i.e., *visual literacy*) and the second lesson focusses on model-based reasoning. The design was based on suggested interventions as described by Quillin and Thomas (2015), and the research was guided by the following two research questions:

> *5.1: What Key activities are developed by the LS-team to foster students' visual literacy and engage students in modelling processes with biological concept-process models when following the suggested interventions by Quillin and Thomas (2015)?*
>
> *5.2: What is the influence of the developed key activities on students' reasoning with biological concept-process models?*

Both developed lessons consisted of multiple key activities, e.g., giving meaning to colours and arrows and creating a model of a biological process (RQ 5.1). Pre- and post-tests and data from student interviews showed that the lessons on visual literacy and model-based reasoning contributed to a more scientific view on the use of concept-process models in science. The lessons improved students' expressed meta-modelling knowledge considering biological concept-process models on multiple levels for both the aspects *nature of models* and *multiple models* (RQ 5.2).

In the studies as described in Chapters 2, 3, 4 and 5, students were only confronted with two-dimensional (2D) biological concept-process models. However, three-dimensional (3D) models in the form of virtual reality (VR) models are starting to find their way into the classroom. In our final study, discussed in **Chapter 6**, we therefore focus on whether students are able to apply the knowledge that they gained while working with 2D concept-process models on a combination of 2D models and a 3D VR model. The following research question was addressed in this study:

> *6.1: Does prior instruction in meta-modelling result in differences in applying meta-modelling knowledge to a combination of 2D models and a 3D virtual environment?*

Our results indicated that the lessons about important aspects of meta-modelling knowledge that the students in the studies described in Chapter 4 and 5 received, did not lead to a preference for two-dimensional models or the virtual reality model when answering questions that are related to important aspects of meta-modelling knowledge. Our results also indicate that students' knowledge about important aspects of meta-modelling knowledge related to 2D models also leads to higher levels of expressed meta-modelling knowledge when working with a combination of 2D models and the VR model. This suggests that students are able to extend the meta-modelling knowledge that they gained while working with 2D models to a 3D VR model.

7.3. Contributions

The main contribution of this dissertation is the evidence found that the developed teaching and learning activities support students' expressed meta-modelling knowledge considering biological concept-process models. Our research project has shed light on the way these lessons can be designed and what the effect of designing these lessons was on both the teachers who were part of the design team (LS-team) and the students. We also developed a test to assess students' expressed meta-modelling knowledge considering biological concept-process models and a VR model that introduces students to the process of blood-glucose regulation.

The effect of the developed lessons on students' meta-modelling knowledge was assessed for the first lesson and the other two lessons separately, and for the combination of the three lessons. The main contribution to students' expressed meta-modelling knowledge was found when looking at the combined effect of the three lessons. This is in line with literature (e.g., An & Cao, 2014), which suggests that supporting student learning over a longer period of time is necessary to develop meta-cognitive skills, such as model-based reasoning. Our research shows that the combination of introducing students to important aspects of meta-modelling knowledge (Lesson 1), stimulating their visual literacy (Lesson 2) and supporting their model-based reasoning skills (Lesson 3) is a fruitful way to stimulate students' expressed meta-modelling knowledge considering biological concept-process models. We have also shown that students' gained meta-modelling knowledge is not limited to the 2D models they worked with during the lessons, but that this knowledge can be extended to a 3D VR model. Since the contribution was mainly visible after all three lessons were taught, our research underlines the importance of making meta-modelling knowledge a recurring theme in the curriculum.

A second contribution concerns the use and evaluation of LS as a research approach. While LS is originally described as an approach that focuses on teachers' professional development (de Vries et al., 2016; Dudley, 2015; Fernandez & Yoshida, 2008), we adapted the approach by adding researchers to the team that designs the lessons (LS-team) and explored the merits of LS as a research approach for research in science education. We found that LS can successfully be used as a research approach to design lessons that answer questions coming from the educational research community, while teachers work on their professional development. We believe that the use of LS as a research approach is a promising option to bring together the pedagogical and didactical knowledge and experience from teachers, and the theoretical knowledge from the educational research community. This exchange in knowledge between teachers and researchers might thereby contribute to bridging the gap between theory-driven research and educational practice. Figure 7.1 shows the adapted version of LS that we used in our research, depicting the roles of the researchers and the teachers in both of the LS-cycles. Figure 7.2 summarizes how LS was used in our research project to determine the effect on students' expressed meta-modelling knowledge. The different steps within a LS-cycle and their connection to the designed key activities is shown. Also, the various sources of data that result from the complete intervention and that can be used to determine students' expressed meta-modelling knowledge are listed.

A third contribution of our research is the instrument we developed to measure students' meta-modelling knowledge considering biological concept-process models for the aspects *nature of models* and *multiple models*. The aim of the developed test was to find out whether students were able to reason on all three levels for these aspects as described in Table 7.1. Tests that use open-ended questions or multiple-choice questions might only measure students' preferred level of reasoning. We therefore used statements, which students had to agree or disagree with, that were related to all three levels for both aspects. This way we were able to describe students' expressed meta-modelling knowledge for each aspect and level separately. This test can be used by other researchers when determining students' expressed meta-modelling knowledge considering biological concept-process models.

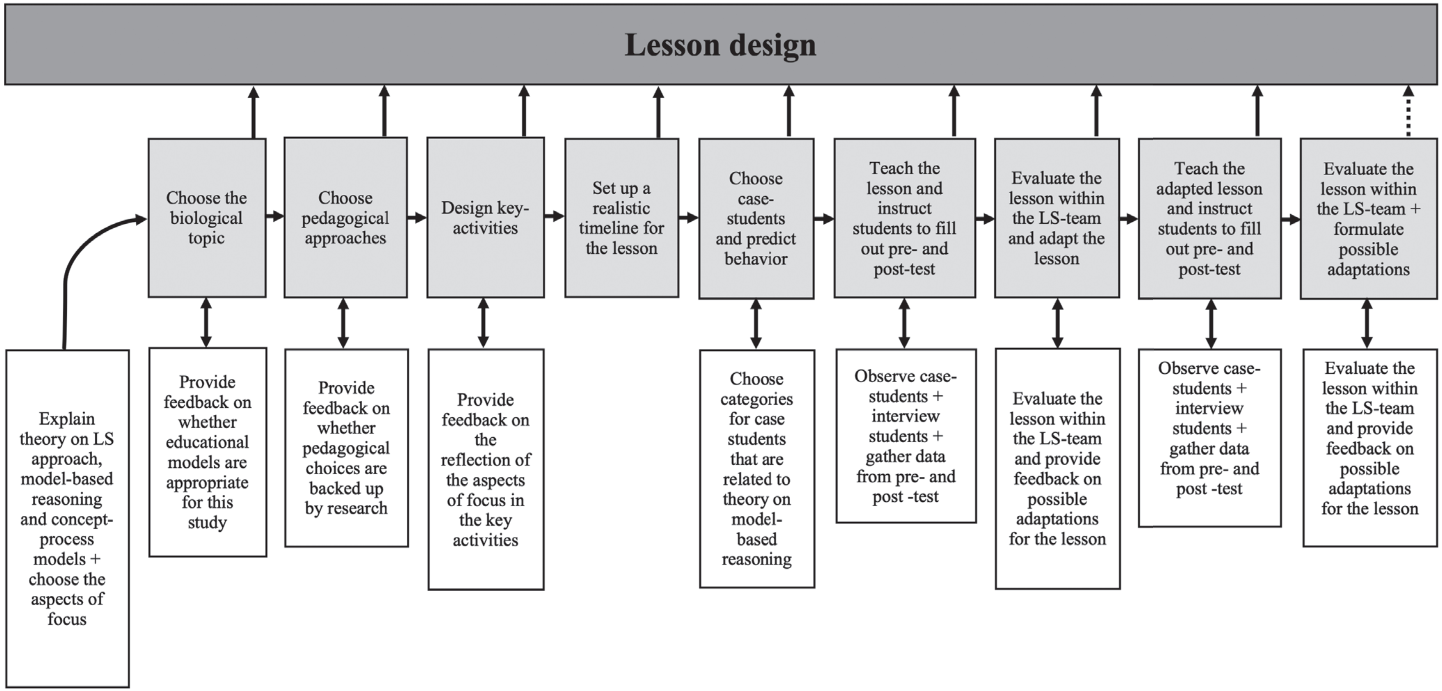


Figure 7.1. Contributions of the teachers and the researchers to the lesson design. The contributions from the teachers are visualized in light grey. The contributions from the researchers are visualized in white. The arrows show interactions between the researchers, teachers and the lesson design. The dotted arrow shows a possible adaptation moment to the lesson design. In our research project these possible adaptations have been discussed, but have not been applied to the lesson design since the lesson would not be taught a third time. The figure can be read as a timeline from left to right.

A fourth and final contribution is the VR model on blood-glucose regulation that we developed. This model was developed with two goals in mind: to introduce students to the process of blood-glucose regulation and to see whether they could apply their gained meta-modelling knowledge from 2D models on a three-dimensional VR model. We mostly see possibilities for further use of the VR model in the classroom, since the model can be used by other teachers to introduce students to the process of blood-glucose regulation. Results showed that students were enthusiastic about the use of VR in class. They indicated that they felt completely emerged into the environment and could work at their own pace. Also, the teachers mentioned that they appreciated the fact that they received student results from the VR model. That way the teachers knew which students needed more help in understanding this biological process.

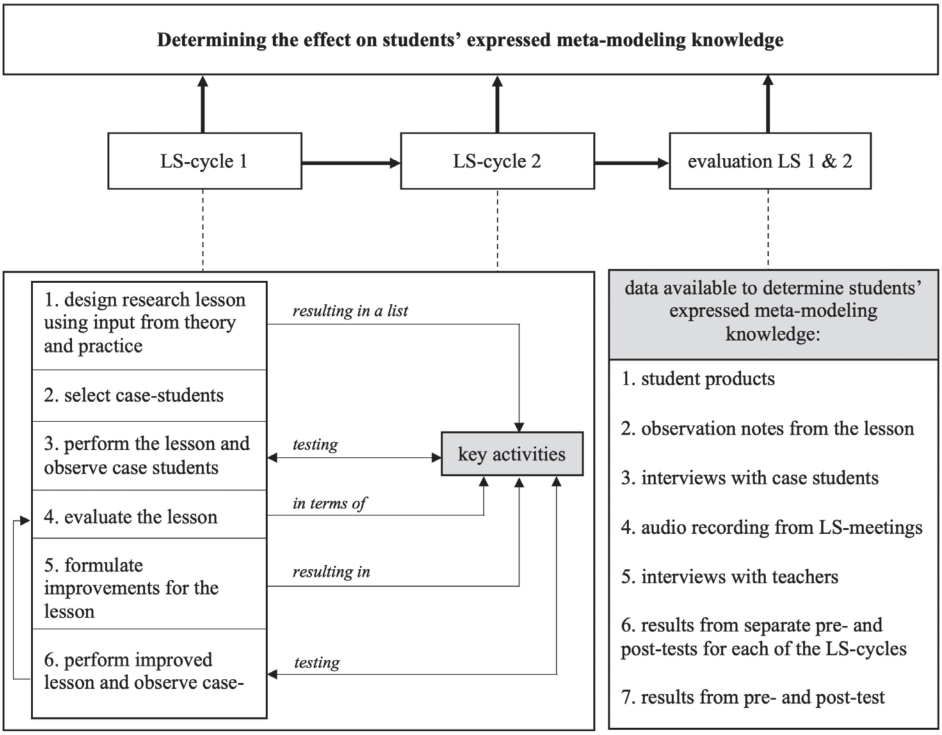


Figure 7.2. Overview of the way data from both LS-cycles and the evaluation of these cycles was combined to determine the effect of both LS-cycles and designed lessons on students' expressed meta-modelling knowledge considering biological concept-process models. The different steps within a LS-cycle and their connection to the designed key activities is shown. Also, the various sources of data that result from the complete intervention and that can be used to determine students' expressed meta-modelling knowledge are listed.

7.4 Limitations and suggestions for future research

The main research question of this research project concerned the support of meta-modelling knowledge considering biological concept-process models. One might wonder whether the results related to students' improved meta-modelling knowledge can fully be attributed to the designed lessons in this research project, since there is no 'control group' in the studies described in Chapter 4 and 5 to compare our results with. It is therefore unclear how students' meta-modelling knowledge develops without receiving these lessons. However, the study described in Chapter 6 does include such a control group and indicates a difference in students' meta-modelling knowledge between the group of students who received the developed lessons as described in Chapter 4 and 5 and the control group who did not receive these lessons. This suggests that the developed lessons did influence students' meta-modelling knowledge. Still, future research should determine the precise attribution of the lessons as described in Chapter 4 and 5 to students' meta-modelling knowledge.

A second limitation of this research project is that the use of LS as described in Chapter 4 and 5 has to been seen as an important part of the way the designed lessons were taught in class. The teachers from the LS-team developed their view on the use of models in education during the LS-cycles and adjusted their teaching accordingly. This means that the lessons as developed by the LS-team might have a different impact on students when teachers who did not join the LS-team teach these lessons. Even though this limitation holds for all lessons that are developed in co-design, three suggestions for future research related to the use of LS can be formulated.

First, it will be interesting to find out what the effect of the developed lessons is on students' meta-modelling knowledge when these lessons are taught by a teacher who was not part of the LS-team. This will give more insight into the effect of teachers being part of the LS-team on students' expressed meta-modelling knowledge.

The second suggestion is related to the first one. If indeed the teachers being part of the LS-team is an important factor for the success of our study, the results as we have seen in this research project cannot easily be transferred to other teachers or schools. Even though we advocate the use of LS as a research approach as an option to sustainably incorporate knowledge from the educational research community into the classroom, we know that this way of using LS cannot be seen as the holy grail for bridging the gap between theoretical knowledge from the educational research community and educational practice. Using LS as a research approach takes a lot of time for the teachers who are part of the LS-team. This means that, in our current education system, it is not feasible to address multiple aspects using LS per year, or have teachers join multiple LS-cycles per year. It might be interesting for future research to find ways in which these results can be transferred. An example could be that teachers who initially joined a LS-team with researchers, afterwards spread their gained knowledge to other teachers in their school and other schools by forming LS-teams with these teachers.

A third option for further research related to the use of LS is to find out what the long-term effects of the three developed lessons is on both teachers and students. The teachers in the LS-team voluntarily incorporated the theory on meta-modelling knowledge in other lessons than the three lessons they designed for this research project. This means that the teachers found the theory important enough to spread to more students than the ones participating in our research. But how long after our research has ended will the teachers keep on incorporating this theory into their lessons? In other words, how sustainable is the implementation of this theory? The same holds for students, of whom we have seen that their meta-modelling knowledge had attenuated after the summer holidays (difference between post-test Chapter 5 and pre-test Chapter 6). As seen in the post-test in Chapter 6, working with the different aspects of meta-modelling knowledge seemed to have brought back this knowledge. But how often do students have to work with these aspects to sustainably support their meta-modelling knowledge? Answering these questions considering the long-term effect on teachers and students would give more insight into how supporting students' meta-modelling knowledge should be incorporated into the curriculum.

A third and fourth limitation of this research project concerns the test that we developed to measure students' expressed meta-modelling knowledge. First of all, the test focuses on only two of the five aspects as described in Table 7.1. This means that the results from the test do not present a complete picture of students' expressed meta-modelling knowledge. Since including all five aspects would lead to too many tasks in the test, the focus on the aspects *nature of models* and *multiple models* seemed the most logical choice for our research project. These aspects mainly focus on reasoning with existing models, which are the models students are reasoning with in the test, in their text books and on exams in school. However, it would be interesting to know the effect of the developed lessons on all five aspects, especially since the third lesson that the LS-team developed involves model-based practices such as testing and changing a model. Future research should therefore focus on developing a test that involves all five aspects of meta-modelling knowledge.

As a fourth limitation, one might wonder whether just ticking a 'yes' or 'no' box is a valid way of measuring students' expressed meta-modelling knowledge. Using interviews with multiple students we did check students' reasoning behind the choices they made, but there is no way to know for every single student why they agreed or disagreed with a statement in the test. Open-ended questions would give more insight in students' way of reasoning, but are not a feasible option when assessing large groups of students. Future research can focus on further establishing the validity of the current test or on developing a test that gives more insight into students' thinking processes. An option could be to replace the 'yes' and 'no' boxes with a Likert scale, as done by Treagust et al., (2002). That way students' answers are not binary (yes/no), but give an indication on how much a student agrees or disagrees with a statement.

As a fifth and final limitation one could argue that our research focuses only on biological concept-process models. Models are used in all STEM courses, which means that meta-modelling knowledge should be stimulated for all types of models in all STEM courses. Since we only focused on biological concept-process models it would be interesting to find out whether students' gained meta-modelling knowledge is transferred to other STEM courses, and if so, whether students apply this meta-modelling knowledge to all types of models, or only to models of processes and phenomena.

7.5 Recommendations for educational practice

In the previous section we outlined some focal points for future research. Still, based on the results of our project, we can formulate recommendations for supporting students' meta-modelling knowledge considering biological concept-process models. In order to support students' meta-modelling knowledge, it is important to incorporate model-based practices into the classroom. We therefore formulate the following recommendations.

Our first recommendation considers the sequence in which aspects that are important for stimulating meta-modelling knowledge should be addressed. Our research shows that the sequence as proposed by Quillin and Thomas (2015), where visual literacy is addressed before students practice with model-based reasoning, supports students' expressed meta-modelling knowledge. However, to give students a framework that they can relate the theory on visual literacy and model-based reasoning to and introduce them to language that is associated with this framework, we started our lesson series with the introduction of three important aspects of meta-modelling knowledge (*nature of models, multiple models* and *purpose of models*). Three Key activities were developed for this lesson that specifically focus on these three aspects (naming differences between models (*multiple models*), matching aims of a model to different models (*purpose of models*), and formulating choices that the creator of a model had made in order to meet the aim of that model (*nature of models*)). For the second and third lesson of the lesson series, we developed Key activities that were in line with the interventions as suggested by Quillin and Thomas (2015), focusing on *visual literacy* and model-based reasoning. All five aspects of meta-modelling knowledge were discussed in these Key activities (*nature of models, multiple models, purpose of models, testing models* and *changing models*). Results from student interviews show that after receiving these three lessons, students use language that is in line with the framework containing descriptions of these five important aspects of meta-modelling knowledge. Therefore, in order to create awareness of these aspects of meta-modelling knowledge and support the use of language that is associated with model-based reasoning, we recommend to not only follow the suggested interventions by Quillin and Thomas (2015), but also to familiarize students with the five aspects of meta-modelling knowledge and the language that is associated with these aspects (Table 7.1).

Second, teachers need to be aware of the importance of meta-modelling knowledge and often need practice in aspects of model-based reasoning themselves. As we have seen in our research project, the theory on different types of models, visual literacy and model-based reasoning can be completely new to teachers. It is unrealistic to expect teachers to educate their students about these aspects when they themselves do not have enough knowledge and experience on this topic. Therefore, it is important to bring teachers into contact with this knowledge. This can be done in various ways, such as implementing these aspects into teacher training courses (pre- and in-service), or, as in our research project, letting teachers be part of a LS-team in which they learn about model-based reasoning and implement this theory in their developed lessons.

Third, once teachers gained enough knowledge on models, visual literacy and model-based reasoning, it is important to make sure that they incorporate model-based practices into the classroom on a regular basis. Model-based reasoning is a meta-cognitive skill, and in order to master such skills practice on a regular basis is necessary (Merchant, Goetz, Cifuentes, Keeney-Kennicutt, & Davis, 2014). However, even though the importance of model-based practices is often described in curricula, most textbooks do not contain any tools that can be used by teachers and students to actually practice with this skill. Of course, it would be best if textbooks incorporated these kinds of practices, but, as shown in our research project, teachers can also develop such activities themselves.

Overall, this dissertation provides an account of a journey from investigating the relatively abstract concept of meta-modelling knowledge in biology domains, exploring its relevance and current status, both among students and teachers to the very concrete design of lessons and lesson materials that can support its development. As such, it can provide an exemplar, for future researchers, developers of educational materials and teacher alike on how this concept – and potentially other abstract concepts – can be taken from theoretical grounds to actual educational practice. As the journey ends here for this dissertation, it can be an inspiration for deeper and more extended explorations in the future.

REFERENCES

References

Ackerman, E., Gatewood, L. C., Rosevear, J. W., & Molnar, G. D. (1965). Model studies of blood-glucose regulation. *The Bulletin of Mathematical Biophysics, 27,* 21-37. https://doi.org/10.1007/BF02477259

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, *16*(3), 183-198. https://doi.org/10.1016/j.learninstruc.2006.03.001

Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to learn in science. *Science, 333*(6046), 1096-1097.

Al-Balushi, S. M. (2011). Students' evaluation of the credibility of scientific models that represent natural entities and phenomena. *International Journal of Science and Mathematics Education, 9*(3), 571–601. https://doi.org/10.1007/s10763-010-9209-4

An, Y.-J., & Cao, L. (2014). Examining the effects of metacognitive scaffolding on students' design problem solving and metacognitive skills in an online environment. *Journal of Online Learning & Teaching, 10*(4), 552-568.

Bakker, A. (2018). *Design research in education: a practical guide for early career researchers* (1st ed.). Routledge.

Bierema, A. M. K., Schwarz, C. V., & Stoltzfus, J. R. (2017). Engaging undergraduate biology students in scientific modelling: analysis of group interactions, sense-making, and justification. *CBE Life Sciences Education, 16*(4). https://doi.org/10.1187/cbe.17-01-0023

de Boer, G. E. (2000). Scientific Literacy: another look at its historical and contemporary meanings and its relationship to science education reform. *Journal of Research in Science Teaching, 37*(6), 582–601.

Bos, A., Marianne, G., Jansen, A., Kalverda, O., de Rouw, T., Smits, G., ... Westra, R. (2012). *Biologie voor jou 5a VWO [Biology for you 5a pre-university level]*. (5th ed.). Malmberg.

Boulter, C., & Buckley, B. (2000). Constructing a typology of models for science education. In J. Gilbert & C. Boulter (Eds.), *Developing models in science education* (pp. 41-57). Kluwer.

Brouwens, R., de Groot, P., & Kranendonk, W. (2013). *BiNaS* (6th ed.). Noordhoff Uitgevers.

Chichibu, T. and Kihara, T. (2013) How Japanese schools build a professional learning community by lesson study, *International Journal for Lesson and Learning Studies*, *2*(1), 12–25. https://doi.org/10.1108/20468251311290105

Cook, M. (2006) Visual representations in science education: the influence of prior knowledge and cognitive load theory on instructional design principles, *Science Education*, *90*(6), 1073-1091. https://doi.org/10.1002/sce.20164

Crawford, B., & Cullin, M. (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. https://doi.org/10.1007/1-4020-3673-6_25

CvTE (2018) *Biologie VWO syllabus centraal examen [Biology pre-university level syllabus for central examination]*. https://www.examenblad.nl/examenstof/syllabus-2020-biologie-vwo/2020/f=/biologie_2_versie_vwo_2020.pdf

Dauer, J. T., Momsen, J. L., Speth, E. B., Makohon-Moore, S. C., & Long, T. M. (2013). Analyzing change in students' gene-to-evolution models in college-level introductory biology. *Journal of Research in Science Teaching, 50*(6), 639-659. https://doi.org/10.1002/tea.21094

Dudley, P. (2015) *Lesson Study, Professional learning for our time*, Routledge, Abingdon.

Dudley, P., Xu, H., Vermunt, J.D., Lang, J. (2019) Empirical evidence of the impact of lesson study on students' achievement, teachers' professional learning and on institutional and system evolution, *European Journal of Education*, *54*(2), 202–217. https://doi.org/10.1111/ejed.12337

Duschl, R. A., Richard A. (1990). *Restructuring science education: the importance of theories and their development*. Teachers College Press.

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

Feinstein, N. (2011). Salvaging science literacy. *Science Education*, *95*(1), 168–185. https://doi.org/10.1002/sce.20414

Fernandez, C., & Yoshida, M. (2004). *Lesson Study: a Japanese approach to improving mathematics teaching and learning*. Lawrence Erlbaum Associates Inc.

Fretz, E. B., Wu, H. K., Zhang, B. H., Davis, E. A., Krajcik, J. S., & Soloway, E. (2002). An investigation of software scaffolds supporting modelling practices. *Research in Science Education, 32,* 567-589. https://doi.org/10.1023/A:1022400817926

Fischer, K., D. Bullock, E. Rotenberg, and P. Raya. 1993. The dynamics of competence. In R. Wozniak and K. Fischer (Eds.), *Development in context* (pp. 93–117). Erlbaum.

Fortus, D., Shwartz, Y., & Rosenfeld, S. (2016). High school students' meta-modelling Knowledge. *Research in Science Education*, *46*(6), 787–810. https://doi.org/10.1007/s11165-015-9480-z

Francoeur, E. (1997). The forgotten tool: the design and use of molecular models. *Social Studies of Science*, *27*, 7–40. https://doi.org/10.1177/030631297027001002

Frigg, R., & Hartmann, S. (2006). Scientific models. In S. Sarkar & J. Pfeifer (Eds.), *The philosophy of science* (pp. 740-749). Routledge.

García-Carmona, A., & Acevedo-Díaz, J. A. (2018). The nature of scientific practice and science education: rationale of a set of essential pedagogical principles. *Science and Education* *27*(5-6), 435-455. https://doi.org/10.1007/s11191-018-9984-9

Gouvea, J., & Passmore, C. (2017). 'Models of' versus 'models for': toward an agent-based conception of modelling in the science classroom. *Science and Education*, *26*, 49–63. https://doi.org/10.1007/s11191-017-9884-4

Gijlers, H., Weinberger, A., van Dijk, A. M., Bollen, L., & van Joolingen, W. (2013). Collaborative drawing on a shared digital canvas in elementary science education: the effects of script and task awareness support. *International Journal of Computer-Supported Collaborative Learning 8*(4), 427-453. https://doi.org/10.1007/s11412-013-9180-5

Gilbert, S. W. (1991). Model building and a definition of science. *Journal of Research in Science Teaching*, *28*, 73–79. https://doi.org/10.1002/tea.3660280107

Gilbert, J., Boulter, C., & Elmer, R. (2000). Positioning models in science education and in design and technology education. In J. Gilbert & C. Boulter (Eds.), *Developing models in science education* (pp. 3-17). Kluwer.

Gilbert, J. K., & Justi, R. (2016). *Modelling-based teaching in science education* (Vol. 9). Springer. https://doi.org/10.1007/978-3-319-29039-3

Gobert, J. D., & Buckley, C. (2000). Introduction to model-based teaching and learning in science education. *International Journal of Science Education, 22*(9), 891-894. https://doi.org/10.1080/095006900416839

Gobert, J. D., O'Dwyer, L., Horwitz, P., Buckley, B. C., Levy, S. T., & Wilensky, U. (2011). Examining the relationship between students' understanding of the nature of models and conceptual learning in biology, physics, and chemistry. *International Journal of Science Education*, *33*(5), 653–684. https://doi.org/10.1080/09500691003720671

Gogolin, S., & Krüger, D. (2018). Students' understanding of the nature and purpose of models. *Journal of Research in Science Teaching*, *55*(9), 1313-1338. https://doi.org/doi:10.1002/tea.21453

Gormally, C., Brickman, P., Hallar, B. and Armstrong, N. (2009) Effects of Inquiry-based Learning on Students' Science Literacy Skills and Confidence, *International Journal for the Scholarship of Teaching and Learning*, *3*(2), Article 16. https://doi.org/10.20429/ijsotl.2009.030216

GOV.UK (2015, May) *National curriculum in England: science programmes of study*, Department for Education. https://www.gov.uk/government/publications/national-curriculum-in-england-science-programmes-of-study/national-curriculum-in-england-science-programmes-of-study

Grosslight, L., Unger, C., Jay, E., & Smith, C. L. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching, 28*(9), 799–822. https://doi.org/10.1002/tea.3660280907

Grünkorn, J., Upmeier zu Belzen, A., & Krüger, D. (2014). Assessing students' understandings of biological models and their use in science to evaluate a theoretical framework. *International Journal of Science Education, 36*(10), 1651–1684 http://dx.doi.org/10.1080/09500693.2013.873155

Harrison, A. G., & Treagust, D. F. (2000). A typology of school science models. *International Journal of Science Education, 22*(9), 1011–1026. https://doi.org/10.1080/095006900416884

Heijnes, D., van Joolingen, W., & Leenaars, F. (2018). Stimulating scientific reasoning with drawing-based modelling. *Journal of Science Education and Technology, 27*, 45-56. https://doi.org/10.1007/s10956-017-9707-z

Hmelo-Silver, C. E., Marathe, S., & Liu, L. (2007). Fish swim, rocks sit, and lungs breathe: expert-novice understanding of complex systems. *Journal of the Learning Sciences, 16*(3), 307-331. https://doi.org/10.1080/10508400701413401

Hodson, D. (2003). Time for action: Science education for an alternative future. *International Journal of Science Education, 25*(6), 645-670. https://doi.org/10.1080/09500690305021

Hodson, D. (2014). Learning science, learning about science, doing science: different goals demand different learning methods. *International Journal of Science Education, 36*(15), 2534-2553.

Hoskinson, A.M., Couch, B. A., Zwickl, B. M., Hinko, K. A., & Caballero, M. D. (2014). Bridging physics and biology teaching through modelling. *American Journal of Physics, 32,* 434. https://doi.org/10.1119/1.4870502

Izquierdo-Aymerich, M., & Adúriz-Bravo, A. (2003). Epistemological foundations of school science. *Science & Education, 12,* 27-43. doi:10.1023/A:1022698205904

Jansen, S., Knippels, M. C. P. J., & van Joolingen, W. R. (2019). Assessing students' understanding of models of biological processes: a revised framework. *International Journal of Science Education, 41*(8), 981-994. https://doi.org/10.1080/09500693.2019.1582821

Jansen, S., Knippels, M.C.P.J., & van Joolingen, W.R. (2021a). Lesson Study as a research approach: a case-study. *International Journal for Lesson and Learning Studies, 10*(3), 286-301.

Jansen, S., Knippels, M.C. P. J., & van Joolingen, W. R. (2021b) *Fostering students' meta-modelling knowledge about biological concept-process models* [Manuscript submitted for publication]. Freudenthal Institute, Utrecht University.

Jansen, S., Knippels, M.C.P.J., van Montfort, M., & van Joolingen, W.R. (2021c). *Secondary school students' meta-modelling knowledge of biological concept-process models* [Manuscript submitted for publication]. Freudenthal Institute, Utrecht University.

Johnstone, A. H. (1991). Why is science difficult to learn? Things are seldom what they seem. *Journal of Computer Assisted Learning, 7*(2), 75-83. https://doi.org/10.1111/j.1365-2729.1991.tb00230.x

Justi, R. S., & Gilbert, J. K. (2002). Science teachers' knowledge about and attitudes towards the use of models and modelling in learning science. *International Journal of Science Education*, *24*(12), 1273–1292. https://doi.org/10.1080/09500690210163198

Justi, R. S., & Gilbert, J. K. (2003). Teachers' views on the nature of models. *International Journal of Science Education, 25*(11), 1369–1386. https://doi.org/10.1080/0950069032000070324

Khishfe, R. (2017). Consistency of nature of science views across scientific and socio-scientific contexts. *International Journal of Science Education*, *39*(4), 403–432. https://doi.org/10.1080/09500693.2017.1287976

King, P. M., & Kitchener, K. S. (2004). Reflective judgment: theory and research on the development of epistemic assumptions through adulthood. *Educational Psychologist*, *39*(1), 5–18. https://doi.org/10.1207/s15326985ep3901_2

Knippels, M.C.P.J., Waarlo, A.J. (2018). Development, uptake, and wider applicability of the yo-yo strategy in biology education research: a reappraisal. *Education Sciences*, *8*, 129. https://doi.org/10.3390/educsci8030129

Kouw, M. (2015). Model-based reasoning in science and technology: theoretical and cognitive issues. *International Studies in the Philosophy of Science, 29*(1), 105-108. https://doi.org/10.1080/02698595.2015.1071555

Kragten, M., Admiraal, W., & Rijlaarsdam, G. (2015). Students' learning activities while studying biological process diagrams. *International Journal of Science Education 37*(12), 1915-1937. https://doi.org/10.1080/09500693.2015.1057775

Krell, M. (2019). Assessment of meta-modelling knowledge: Learning from triadic concepts of models in the philosophy of science. *Science Education Review Letters*. Online publication. https://doi.org/10.18452/19813

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. https://doi.org/10.1080/00219266.2015.1028570

Krell, M., & Krüger, D. (2017). University students' meta-modelling knowledge. *Research in Science and Technological Education*, *35*, 261–273. doi:10.1080/02635143.2016.1274724.

Krell, M., Reinisch, B., & Krüger, D. (2015). Analyzing students' understanding of models and modelling referring to the disciplines biology, chemistry, and physics. *Research in Science Education, 45*(3), 367–393. https://doi.org/10.1007/s11165-014-9427-9

Krell, M., Upmeier zu Belzen, A., & Krüger, D. (2012). Students' understanding of the purpose of models in different biological contexts. *International Journal of Biology Education, 2*(2), 1–34. https://doi.org/10.20876/IJOBED.94733

Krell, M., Upmeier zu Belzen, A. & Krüger, D. (2014). Context-specificities in students' understanding of models and modelling: An issue of critical importance for both assessment and teaching. 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* (pp. 1024-1035). European Science Education Research Association.

Lee, K. (2012). Augmented reality in education and training. *TechTrends, 56*, 13-21. https://doi.org/10.1007/s11528-012-0559-3

Lehrer, R., & Schauble, L. (2006). Scientific thinking and science literacy. In K.A. Renninger, I.E. Sigel, W. Damon, & R.M. Lerner (Eds.), *Handbook of Child Psychology* (pp. 153–196). John Wiley & Sons.

Louca, L. T., Zacharia, Z. C., & Constantinou, C. P. (2011). In quest of productive modelling-based learning discourse in elementary school science. *Journal of Research in Science Teaching, 48*(8), 919–951. https://doi.org/10.1002/tea.20435

Louca, T.L., Zacharia, Z., Michael, M., & Constantinou, P.C. (2011). Objects, entities, bahaviors and interactions: a typology of student-constructed computer-based models of physical phenomena. *Journal of Educational Computing Research, 44,* 173-201. doi:10.2190/EC.44.2.c

Lundström, M., Ekborg, M., & Ideland, M. (2012). To vaccinate or not to vaccinate: How teenagers justified their decision. *Cultural Studies of Science Education, 7*, 193-221. https://doi.org/10.1007/s11422-012-9384-4

Mahr, B. (2009). Information science and the logic of models. *Software and Systems Modelling*, *8*, 365–383. https://doi.org/10.1007/s10270-009-0119-2

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers and Education, 70*, 29-40. https://doi.org/10.1016/j.compedu.2013.07.033

van Meter, P., Aleksic, M., Schwartz, A., & Garner, J. (2006). Learner-generated drawing as a strategy for learning from content area text. *Contemporary Educational Psychology, 31*(2), 142-146. https://doi.org/10.1016/j.cedpsych.2005.04.001

Ming Cheung, W. and Yee Wong, W. (2014) Does Lesson Study work?: A systematic review on the effects of Lesson Study and Learning Study on teachers and students. *International Journal for Lesson and Learning Studies*, *3*(2), 137-149. https://doi.org/10.1108/IJLLS-05-2013-0024

Miri, B., David, B.C. and Uri, Z. (2007) Purposely teaching for the promotion of higher-order thinking skills: a case of critical thinking. *Research in Science Education*, *37*, 353-369. https://doi.org/10.1007/s11165-006-9029-2

van Montfort, M. (2019). *Secondary school students' meta-modelling knowledge of concept-process models in biology* (unpublished master's thesis). Utrecht University, Utrecht.

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

NGSS Lead States. (2013). *Next Generation Science Standards: For states, by states*. The National Academies Press. https://doi.org/10.17226/18290

Novick, L. R., & Catley, K. M. (2007). Understanding phylogenies in biology: the Influence of a gestalt perceptual principle. *Journal of Experimental Psychology: Applied, 13*(4), 197-223. https://doi.org/10.1037/1076-898X.13.4.197

Odenbaugh, J. (2005). Idealized, inaccurate but successful: A pragmatic approach to evaluating models in theoretical ecology. *Biology and Philosophy, 20*, 231–255. https://doi.org/10.1007/s10539-004-0478-6

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. https://doi.org/10.1080/09500693.2010.502191

Opfer, V.D. and Pedder, D. (2011) Conceptualizing teacher professional learning. *Review of Educational Research*, *81*(3), 376–407. https://doi.org/10.3102/0034654311413609

Passmore, C., Gouvea, J. S., & Giere, R. (2014). Models in science and in learning science: Focusing scientific practice on sense-making. In M.R. Mathews (Ed.), *International handbook of research in history, philosophy and science teaching* (pp. 1171–1202). Springer. https://doi.org/10.1007/978-94-007-7654-8_36

Quillin, K., & Thomas, S. (2015). Drawing-to-learn: A framework for using drawings to promote model-based reasoning in biology. *CBE-Life Sciences Education, 14*(1), 1-16. https://doi.org/10.1187/cbe.14-08-0128

Roberts, R., & Gott, R. (2010). Questioning the evidence for a claim in a socio-scientific issue: An aspect of scientific literacy. *Research in Science and Technological Education, 28*(3), 203-226. https://doi.org/10.1080/02635143.2010.506413

Schipper, T., Goei, S.L., de Vries, S. and van Veen, K., (2017) Professional growth in adaptive teaching competence as a result of Lesson Study. *Teaching and Teacher Education*, *68*, 289-303. https://doi.org/10.1016/j.tate.2017.09.015

Schönborn, K. J., & Anderson, T. R. (2009). A model of factors determining students' ability to interpret external representations in biochemistry. *International Journal of Science Education*, *31*(2), 193-232. https://doi.org/10.1080/09500690701670535

Schwamborn, A., Mayer, R. E., Thillmann, H., Leopold, C., & Leutner, D. (2010). Drawing as a generative activity and drawing as a prognostic activity. *Journal of Educational Psychology, 102*(4), 872-879. https://doi.org/10.1037/a0019640

Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., … Krajcik, J. (2009). Developing a learning progression for scientific modelling: Making scientific modelling accessible and meaningful for learners. *Journal of Research in Science Teaching, 46*(6), 632-654. https://doi.org/10.1002/tea.20311

Schwarz, C. V., & White, B. Y. (2005). Meta-modelling knowledge: developing students' understanding of scientific modelling. *Cognition and Instruction*, *23*, 165–205. https://doi:10.1207/s1532690xci2302_1

Sharon, A. J., & Baram-Tsabari, A. (2020). Can science literacy help individuals identify misinformation in everyday life? *Science Education*, Article sce.21581. https://doi.org/10.1002/sce.21581

Shim, K. C., Park, J. S., Kim, H. S., Kim, J. H., Park, Y. C., & Ryu, H. Il. (2003). Application of virtual reality technology in biology education. *Journal of Biological Education, 37*(2), 71-74. https://doi.org/10.1080/00219266.2003.9655854

Smith, M.K., Wood, W.B., Adams, W.K., Wieman, C., Knight, J.K., Guild, N. and Su, T.T., (2009). Why peer discussion improves student performance on in-class concept question. *Science*, *323*(5910), 122-124. http://www.jstor.org/stable/20177137.

Stern, E., Aprea, C., & Ebner, H. G. (2003). Improving cross-content transfer in text processing by means of active graphical representation. *Learning and Instruction, 13*(2), 191-203. https://doi.org/10.1016/S0959-4752(02)00020-8

Svoboda, J., & Passmore, C. (2013). The strategies of modelling in biology education. *Science and Education, 22*(1), 119–142. https://doi.org/10.1007/s11191-011-9425-5

Tanner, K.D. (2009). Talking to learn: Why biology students should be talking in classrooms and how to make it happen. *CBE Life Sciences Education*, *8*, 89-94. https://doi.org/10.1187/cbe.09-03-0021

Tasker, R. and Dalton, R. (2007) Visualizing the molecular world – design, evaluation, and use of animations. In Gilbert, J.K., Reiner, M. & Nakhleh, M. (Eds.), *Visualization: Theory and Practice in Science Education* (pp. 103-131). Springer.

Treagust, D. F., Chittleborough, G., & Mamiala, T. L. (2002). Students' understanding of the role of scientific models in learning science. *International Journal of Science Education, 24*(4), 357–368. https://doi.org/10.1080/09500690110066485

Treagust, D. F., Chittleborough, G. D., & Mamiala, T. L. (2004). Students' understanding of the descriptive and predictive nature of teaching models in organic chemistry. *Research in Science Education*, *34*, 1-20 https://doi.org/10.1023/B:RISE.0000020885.41497.ed

Tsui, C. Y., & Treagust, D. F. (2013). Introduction to multiple representations: Their importance in biology and biological education. In Treagust D., Tsui CY. (Eds.), *Multiple representations in biological education*. *Models and modelling in science education* (pp. 3-18). Springer. https://doi.org/10.1007/978-94-007-4192-8_1

Upmeier zu Belzen, A., & Krüger, D. (2010). Modellkompetenz im biologieunterricht [Model competence in biology education]. *Zeitschrift Für Didaktik Der Naturwissenschaften*, *16*, 41–57.

Upmeier zu Belzen, A., van Driel, J., & Krüger, D. (2019). Introducing a framework for modelling competence. In Upmeier zu Belzen, A., van Driel, J., & Krüger, D. (Eds.), *Models and modelling in science education* (Vol. 12) (pp. 3–19). Springer. https://doi.org/10.1007/978-3-030-30255-9_1

Velev, D., & Zlateva, P. (2017). Virtual reality challenges in education and training. *International Journal of Learning and Teaching, 3*(1), 33-37. https://doi.org/10.18178/ijlt.3.1.33-37

Verhoef, N.C. and Tall, D.O. (2011). Lesson study: the effect on teachers' professional development., *35th Conference of the International Group for the Psychology of Mathematics Education (PME)*, 10-15 July 2011, Ankara, Turkey.

Vermunt, J.D., Vrikki, M., van Halem, N., Warwick, P. and Mercer, N., (2019). The impact of Lesson Study professional development on the quality of teacher learning. *Teaching and Teacher Education*, *81*, 61–73. https://doi.org/10.1016/j.tate.2019.02.009

de Vries, S., Verhoef, N. and Goei, S. L. (2016) *Lesson study: een praktische gids voor het onderwijs [Lesson Study: a practical guide for education],* Garant Uitgevers.

White, B. Y., Collins, A., & Frederiksen, J. R. (2011). *The Nature of Scientific Meta-Knowledge*. *Models and Modelling*. Springer. https://doi.org/10.1007/978-94-007-0449-7

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. https://doi.org/10.1002/sce.20259

Wolthuis, F., van Veen, K., de Vries, S. and Hubers, M.D., (2020). Between lethal and local adaptation: Lesson study as an organizational routine. *International Journal of Educational Research*, *100*, 101534. https://doi.org/10.1016/j.ijer.2020.101534

Zhou, F., Dun, H. B. L., & Billinghurst, M. (2008). Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR. *Proceedings of the International Symposium on Mixed and Augmented Reality, UK,* 193-202. https://doi.org/10.1109/ISMAR.2008.4637362

8

APPENDICES

Appendix A

Interview questions as used in the study described in Chapter 2.

Questions related to the model showing accumulation of DDT in the food chain:

1. What is depicted here?
2. What do the arrows mean?
3. Why are the numbers 1 through 4 next to the arrows?
4. What does this picture not show regarding the food chain?
5. Can you comment on the extent to which this image represents reality?
6. What would be the purpose of this image?
7. Could the image be used for another purpose?
8. For what purpose is this image particularly appropriate, and why?
9. How is it determined whether this image is correct?
10. More images on the same theme exist. Why is there not just one image about this theme?
11. The image below is about the same theme. Can you name differences in the information given in both images?
    [Model also showing the accumulation of DDT in the food chain, but with a focus on biomass]
12. Does this process also affect the cellular level? And the population level?
13. What is the reason that the influence from question 12 is not visible in this image? Do you think that images where this influence is visible exist?
14. For both images, name a situation in which the use of that image is appropriate.
15. It could be that the image will be adapted. What would be a good reason to modify the image and what would be the modification?

Questions related to the model showing the process of spermatogenesis:

1. What is depicted here?
2. What do the arrows mean? (just repeating the legend is not enough. For example, what does stimulate mean?)
3. Round shapes, oval shapes, and semicircle shapes are used in the source, among others. Why were these shapes chosen?

4. To what extent are these shapes based on what these substances look like in reality?
5. What does this image not show regarding spermatogenesis?
6. Can you comment on how well this source represents reality?
7. What would be the purpose of this source?
8. Could the image be used for another purpose?
9. For what purpose is this image most appropriate, and why?
10. How is it determined if the source is correct?
11. There do exist more images on the same theme. Why is there not just one image about this theme?
12. The image below is about the same theme. Can you name differences in the information given in both images?
    [Model also showing the process of spermatogenesis, but with a focus on meiosis]
13. Does this process also have an effect on the biological level of organisms?
14. What is the reason that the influence from question 12 is not visible in this image? Do you think that images where this influence is visible exist?
15. For both images, name a situation in which the use of that image is appropriate.
16. It could be that the image will be adapted. What would be a good reason to modify the image and what would be the modification?

Table A. Overview of the relationship between the interview questions and the aspects of meta-modelling knowledge.

Aspects	DDT model	Spermatogenesis model
Introduction of model	1, 2, 3	1, 2, 3
Nature of models	4	4
Multiple models	9, 10, 11	10, 11, 12
Purpose of models	5, 6, 7	5, 6, 7, 8
Testing models	8	9
Changing models	12	13

Appendix B

Framework used to code student answers in the study described in Chapter 2. Each aspect of meta-modelling knowledge is shown in a separate table (Table B1 – B5). The first column shows the aspect of meta-modelling knowledge, along with the three or four levels of meta-modelling knowledge (depending on the aspect), and the description of these levels and aspects as defined by Grünkorn et al. (2014). The descriptions in *red* were added during the study as described in Chapter 2, based on student answers from the interview. The second column shows examples of student answers as mentioned by Grünkorn et al. (2014). The third column shows for the aspects and levels examples of student answers as found in the study described in Chapter 2.

In the study described by Grünkorn et al. (2014), categories were defined within some of the levels and aspects. The first two columns in Table B1 – B5 show these categories in separate rows. In our study, as described in Chapter 2, we only coded the aspects and levels, not the different categories. Therefore, the third column shows examples of student answers for each of the aspects and levels, without distinguishing different categories.

Table B1. Framework to code student answers related to the aspect *nature of models*.

Nature of models	**Examples of student answers as mentioned by Grünkorn et al. (2014)**	**Examples of student answers from our own study (Chapter 2)**
Level I	Level I	Level I
Model as copy **– Matches the original** **– Enlarged/reduced scale copy of original** **– Accepted as scale model of the original, because there is great confidence in science, in the scientific method, or in the scientists**	- ‘The Tyrannosaurus rex looked like the model’ (student Qb366) - ‘The model of a biomembrane is an enlarged copy of a biomembrane’ (student Qb502) - ‘The Neanderthal man looked as it does in the model, because many biologists have certainly worked on this model. These people know what they are talking about, so he looked this way’ (student Qa485)	- I think it reflects reality quite well, as it is shown step by step: look, you start with the smallest thing here, then that goes on like that… realistic steps, then. - We saw something like that in class once. The hypothalamus controlled the whole process and then sperm cells were made - I think this [model] has something to do with food. It shows that the algae use the DDT as a food source. And small fish eat algae. And carnivorous fish eat the smaller fish. And then the heron, a fish-eating bird, eats the other fish.
Model with great similarity **– Resembles the original** **– Nearly scale model of the original due to dissatisfaction with the modelling process**	- ‘The model of the biomembrane is very similar to the real biomembrane. Both have a surface layer that holds it all together and both have tissue in the centre’ (student Qd1073) - ‘The model resembles the Neanderthal man. Only the place on the model where the eyebrows are has to be pushed forward a little, because that’s how it is in the skeletal findings. That has to happen; otherwise, the model would be incorrect’ (student Qd567)	
Model represents a (non-) subjective conception of the original **– Compares and judges the model based on prior knowledge of, personal experience of, or subjective conceptions about the original** *- refers to what has been taught in class about the concepts mentioned in the model*	- ‘I don’t think the model is correct. The real Neanderthal man looked more like an ape. That’s how I imagine one’ (student Qd915)	
Displays a process, its components and how they are related **- describes the process, believing the relationships as shown can be seen as a copy of the original**		

Nature of models	**Examples of student answers as mentioned by Grünkorn et al. (2014)**	**Examples of student answers from our own study (Chapter 2)**
Level II	Level II	Level II
Parts of the model are a copy **– Only certain features resemble the original; other features cannot be judged due to paucity of information or knowledge about the original**	'The skeletal findings and the model have the same head shape. It is unknown whether the hair back then was the way it is now. Nothing can be stated with certainty about the eyes, either. On the whole, one can only comment on the shape. Colour and such remains "unknown"' (student Qa97)	- I think there is a lot more going on in the body, also in those places, but well, that is not important for now, so that is not mentioned.
Model as a possible variant **– Might resemble the original (or not); abstract statements about similar properties – One conceivable version among many, but less well founded**	- 'The model is comparable in terms of its shape. Nonetheless, one cannot assume that the Neanderthal man really looked like this' (student Qa164) - 'Yes, it [the biomembrane] might look like that, but it might also look like this [picture drawn by student]' (student Qb417)	
Model as focused representation **– Focused on one element of the original, highlights certain traits/ properties**	'The model only shows the essential parts of a real biomembrane. The main traits, structures and colours are shown here' (student Qb6)	
Level III	Level III	Level III
· ***Model as hypothetical representation*** **– Presents a justified hypothesis about the original, possible similarity between original and model is discussed**	- 'No one can know for certain what a living Tyrannosaurus rex looked like back then. Scientists can only make assumptions about how it looked. They analyse the skeleton and use that to calculate how its body might have been constructed' (student Qb24)	- It is not possible to see this in real life. [..] Maybe in the computer you could simulate that or something

Table B2. Framework to code student answers related to the aspect *purpose of models.*

Purpose of models	Examples of student answers as mentioned by Grünkorn et al. (2014)	Examples of student answers from our own study (Chapter 2)
Level I	Level I	Level I
***Model for showing the facts* – Presenting the facts**	'The model shows the different plants that grow in a forest' (student Qa403)	- To show how hormones work, and well, do their job in spermatogenesis - I think, well, perhaps for students. To show them like, here you have producers and here you have consumers. So, you can explain that these eat these [pointing at organisms in model], and so on
***Model for showing events* - Showing in what way a process functions**		
Level II	Level II	Level II
***Model to identify relationships* – Describing relationships between different aspects in the original and serving to understand known facts**	'The model shows that it is possible to observe how the leaves and blossoms develop and spread' (student Qa217)	- I think they are trying to explain DDT, if it is in the water, what consequences that will have
***Model to explain relationships* – *Not just* describing*, but also explaining* relationships between different aspects in the original and serving to understand known facts**	'It is meant to demonstrate that the sea is a good habitat for animals e.g. fish and plants. It also shows and explains how the sea constitutes a "circulatory chain". The plants could not survive without the oxygen and the water; the fish could not live without the plants' (student Qb175)	
Level III	Level III	Level III
***Model to examine abstract/ concrete ideas* – Serving as an instrument to test hypotheses about the original; general ideas are mentioned – Both testing hypotheses about the original and serving to draw conclusions about the original; concrete ideas are mentioned – Serving to transfer findings about the original to other phenomena**	- 'One could also test which plants grow best and most quickly in which types of soil and compare these results. It might be that a certain plant draws so many nutrients out of the soil that there are less available for another. If the model were to prove this, we would know that these types of plant should not be planted too closely together' (student Qd263)	- Suppose you want someone to get more or less testosterone in his body, then you can see with which hormones you can inhibit or stimulate it, and you can also choose which part of the body you can best add it to

Table B3. Framework to code student answers related to the aspect *multiple models*.

Multiple models	Examples of student answers as mentioned by Grünkorn et al. (2014)	Examples of student answers from our own study (Chapter 2)
<u>Initial level 0</u>	<u>Initial level 0</u>	<u>Initial level 0</u>
All models are the same **– All models are or show the same; no description of differences between models**	'All three models show basically the same. I don't know why there should be different models at all. That makes no sense' (student Qd240)	- Some people find one image more chill than another. One is clearer than the other and everyone who makes it, I assume, hopes he has the most clear one
Various models of different originals **– Each model represents a different original**	'One might also make three different models to show the biomembranes of different life forms, e.g. a human being, a bird and a cow' (student Qb331)	
Only one final and correct model **– Only one of the various models is final and correct; the others are incorrect** **– Only one model is the final model; they are not valid contemporaneously**	- 'Perhaps there is only model which is the best model. All the others are wrong. There is just one correct model. How else should this work?' (student Qd643) - 'I think that two models are old models. At that time, one did not have all information like we have today. However, one model is the final model which happens to be true. I've seen it in my textbook' (student Qd321)	

Multiple models	Examples of student answers as mentioned by Grünkorn et al. (2014)	Examples of student answers from our own study (Chapter 2)
Level I ***Different model object properties*** **– Differing methods of presentation (2D or 3D, different colours, etc.)** **– Differing model features (moveable or immoveable, soft or hard, large or small, etc.)** **– Differing construction options (thin or thick materials, separated elements or one piece, etc.)**	Level I - 'The tongue can be shown a little differently. No colours are used in Model C, for example, whereas they are in Models A and B' (student Qc62) - 'Models B and C vary in their solidity. Model B is stiffer while Model C is very flexible' (student Qb64) 'The different elements are easy to recognise in Model A. The rods on Model C are thinner than in B' (student Qb311)	Level I - I'm sure there are also models that are much more schematic. You can also use blocks, for example. You could also simply say that this block has a certain process and that these substances are produced within it, and then you could use arrows to indicate that process. So it can be less specific, more schematic
Level II ***Focus on different aspects*** **– The complexity of the original allows diverse perspectives or ways of focusing on the original (interior or exterior, profile or cross-section, structure or function, diverse sections or states of the original, etc.)**	Level II Since each of these models highlights something different, there are different models. Model A focuses on the different elements and the structure, while Model B and C look more at the construction of a biomembrane' (student Qd744)	Level II - Otherwise it is too much information in one picture - Because you can also display this in different ways. For some people, this will be very confusing with these pictures and you may also want to emphasise other things, I think

Multiple models	**Examples of student answers as mentioned by Grünkorn et al. (2014)**	**Examples of student answers from our own study (Chapter 2)**
Level III	Level III	Level III
Different assumptions **– There can be various assumptions and ideas about the original; different models are valid at the same time** **– Differing interpretations of the data** **- Different assumptions with prospects of application** **– Differing assumptions about the original are named after scientific purposes (basis of discussion, comparison of different assumptions, testing assumptions with the models, etc.)**	- 'Since there are various theories/ideas about the human oesophagus, there will also be alternative models. Scientists might have other opinions' (student Qb5) - 'The persons have drawn different conclusions from their observations, which is why there are different models of this biomembrane' (student Qd468)	- I think there are different ideas about what exactly is happening, so you can also visualise it in different ways

Table B4. Framework to code student answers related to the aspect *testing models*.

Testing models	**Examples of student answers as mentioned by Grünkorn et al. (2014)**	**Examples of student answers from our own study (Chapter 2)**
Initial level 0	Initial level 0	Initial level 0
No testing of models **– Rejecting model testing in general or of this model**	'Why should this beetle model even be tested? I don't think it's necessary' (student Qc71)	- I think that the creator of the schoolbook, that he knows how it works
Perceiving schoolbooks or their authors as authorities providing absolute truth **- Comparing the model with knowledge from literature, because there is great confidence in science or in the scientists**		

Testing models	Examples of student answers as mentioned by Grünkorn et al. (2014)	Examples of student answers from our own study (Chapter 2)
Level I	Level I	Level I
Testing of material **– Testing the resistance of the material (flexibility, stability, elasticity, weight, etc.)**	- 'One should test if the material of the model is strong enough to remain unharmed by something such as wind' (student Qd793)	- I assume that they have investigated this in a lab, and that there is very good research behind this
Testing of basic requirements **– Naming fundamental requirements for that model**	- 'For starters, the model should be able to fly in any case. Otherwise, I don't think the model would be very good' (student Qa4)	
Level II	Level II	Level II
Comparison between original and model **– Comparing the properties (structure and/or function) of the original with those of the model**	- 'The model has to be compared to a real beetle' (student Qc69)	- You could do research on it, microscopic or so and then you could see how everything reacts from one substance to another and so on
Comparison and matching of original and model **– Both comparing properties and describing the necessary adjustments for congruity between the model and the original; naming criteria for a good model**	- 'The model can be tested for its dimensions, its weight. The structure of the model must match the original or it isn't suitable' (student Qb206)	

9

Testing models	**Examples of student answers as mentioned by Grünkorn et al. (2014)**	**Examples of student answers from our own study (Chapter 2)**
<u>**Level III**</u> ***Testing hypotheses*** **– Testing hypotheses about the original using the model and listing general ideas for studies**	<u>Level III</u> 'This model could simulate the flight of such a seed. Such simulations would show where the seed flies to and how it gets implanted into the soil. The model could also be used to test the effects the impact has on the soil, on the flight, and on the seed' (student Qb278)	<u>Level III</u> - By doing a lot of research, actually looking at how those food chains are put together and whether there is really a problem if you were to take away the herbivorous fish, for example, step 2 to 3. You can then look what would happen in that case and whether it is true that the whole food chain would be a bit shaken up.
Testing of hypotheses with research designs **– Describing a concrete application for the model (research design) to test a hypothesis about the original**	'One has to try to obtain videos of the original flight manoeuvres and attempt to recreate and compare these with the model in a wind tunnel to see if the model behaves as the original. If so, one has to change the environmental influences in the wind tunnel to determine what the dragonfly needs to fly' (student Qb200)	

Table B5. Framework to code student answers related to the aspect *changing models*.

Changing models	Examples of student answers as mentioned by Grünkorn et al. (2014)	Examples of student answers from our own study (Chapter 2)
Initial level 0	Initial level 0	Initial Level 0
No reason for alterations **– Rejecting changes to a model**	'I don't think the model should be changed' (student Qd341)	- I have no idea
Alteration of how different originals are represented **– Creating different models for different originals; each original is represented by its own model**	'Because not all dragonflies are alike and models can be made for different dragonflies' (student Qd954)	

Changing models	Examples of student answers as mentioned by Grünkorn et al. (2014)	Examples of student answers from our own study (Chapter 2)
Level I	Level I	Level I
Alterations to improve the model object **– Optimising the functioning/aesthetics of the model object** **– Optimising the technology of model Creation**	- ‘The only reason why most models are changed is because their movement and functionality can be improved’ (student Qa378) - ‘To change the model of the dragonfly, a new technology is needed that allows the model to stay up without needing to attach a booster for uplift on the long back legs’ (student Qb521)	- Maybe if they want to make it clearer, like for this one, there are no transport arrows visible. You could add them. So that it is even clearer what processes are taking place. - That they find out that the image is not correct after all
Alterations when there are errors in the model object **– Fundamental considerations for fixing errors in the model** **– Referencing concrete, incorrect properties of the model (e.g. defective materials)**	- ‘I think it’s because errors are always being found which need to be corrected’ (student Qd629) - ‘Perhaps the wings have to be made out of harder materials; otherwise they cannot resist the pressure during flight’ (student Qb508)	
Alterations when basic requirements are not met **– Reviewing the basic requirements of each model and correcting defects if necessary**	‘If the model is meant to fly and it doesn’t, then the scientists have to work on it’ (student Qb496)	

Changing models	Examples of student answers as mentioned by Grünkorn et al. (2014)	Examples of student answers from our own study (Chapter 2)
Level II	Level II	Level II
Alterations when model does not match the original **– Optimising how the (structure and/or function of the) model matches the original with consideration of the necessary congruity between the original and the model**	- 'The model doesn't look exactly like a crab. The legs of a real crab are longer. The body of a crab is somewhat narrower. This is not the shape of a crab. That should definitely be changed, because it has to match the real crab' (student Qd1010)	- New discoveries in science
Alterations due to new findings about the original **– Integrating new findings about the original into the model; improved technology leads to new findings about the original** **Alterations due to changes in the original** **– Reflecting changes (e.g. individual developments) or advancements (e.g. evolutionary adaptation) in the original as new information in the model**	- 'In a few years, we will have better technology, so we can learn more about the dragonfly. The model could be changed when something new about the dragonfly is discovered' (student Qd1199) - 'There are always changes in biology and in history. The same is true of the crab. Evolution changes the environment and animals have to adapt again. Changes to the environment force animals to change as well. That's why the model can be changed' (student Qd1165)	

Changing models	**Examples of student answers as mentioned by Grünkorn et al. (2014)**	**Examples of student answers from our own study (Chapter 2)**
Level III ***Alterations due to findings from model experiments*** **– Adjusting the model to reflect findings about the original based on a model experiment or falsification of the hypothesis behind the model** *Alterations when the focus of the model shifts to a different aspect of the process* *- Adjusting the model to lay the focus on a different part of the process that is depicted*	Level III - 'If tests of a flying object show that the model flies completely differently than thought or than a real dragonfly does, then something could be changed on the gliding surfaces. The scientists may have had a different assumption' (student Qd352)	Level III - When the purpose of the image is different. This is quite abstract I think. So, in theory, well this is mostly about the nucleus. So, you could also make clear how the rest of the cell is being influenced by this process. Or perhaps in which organs this happens

Appendix C

The test used in Chapters 3, 5 and 6 is available online using the following link:
https://survey.uu.nl/jfe/form/SV_8Cvo34d7TwobxAx

Appendix D

List of translated questions from the pre- and post-test from the study as described in Chapter 5. The questions are related to the topics of the developed lessons: visual literacy and model-based reasoning.

1. In a model, symbols such as arrows are often used. The figure below shows three examples. Describe what you think is the general meaning of arrows in a model of a process. Note that this question is about the use of arrows in general, not about the meaning of a specific arrow from one of the models below.

[model showing the way reflexed work in the nervous system, model showing the process of pollination and model showing the way a human body keeps its temperature at 37 degrees Celsius]

2. The text below is copied from a biology textbook. If you wanted to draw a model of this process, you would have to decide which terms from the text you would draw in this model and how to connect these terms using arrows. Which terms would you draw and connect if you were to make a model of the following text?

"Oxytocin stimulates the - reflexive - contractions of the smooth muscle tissue in the mammary glands of the breasts. The baby's sucking on the nipple is a stimulus for the thalamus to produce more oxytocin. The milk expulsion stimulated by this is in addition to the milk flow caused by the baby's own sucking".

3. In the picture below you find two models on the same subject, drawn on a different level of biological organisation. What is the reason that there are several models on the same subject at different levels of biological organisation?

[model showing a heredity family tree and a model showing a pair of chromosomes with specific alleles]

4. In the previous task you saw an example of models of the same process drawn at a different level of biological organisation. However, there are also models of the same process at the same level of biological organisation. In these models, is one of the models always better than the other?

5. Each model has a core message. Please indicate what the core message of the model below is.

[model containing a graph, showing the density of a population over time]

In this last part of the test we will look at how well you can work with models of control loops. Control loops ensure that in your body different values are kept at a certain level. You can think of processes that maintain your blood sugar level or your temperature.

6. For the maintenance of a certain standard value, a 'sensor', a 'control centre' and an 'effector' are important. Explain why these three elements are important for maintaining a certain value at the set level.

7. The word control loop indicates that the process takes place in a loop. In what way are the three elements connected?

Explain why you have chosen this option.

SUMMARY

Summary

Understanding scientific information is not only important for scientists, but also for non-scientists. Models showing data about recent scientific discoveries are shown on the news, in newspapers and on social media. An example of how this transfer of information into society happens can be seen when looking at the recent COVID-19 pandemic outbreak (2020). Models are popping up everywhere, showing how herd immunity can protect people, how vaccination works, or what the prognosis considering the course of the pandemic is. This transfer of scientific information in the form of models into society makes it important for citizens to understand the creation and use of models in science. Education can play an important role in stimulating students' knowledge about models, the creation of models and the use of models (i.e., meta-modelling knowledge). This is why many countries adopted modelling as a scientific practice into their curricula. However, the literature on this topic shows that in practice teachers mainly use models as an aid in teaching science content, neglecting the scientific practices behind these models. Students are therefore often not explicitly taught the meta-modelling knowledge that is required to understand the use of models in science. Absence of knowledge about for example the fact that models are often simplifications of reality and only display what is currently known or hypothesized about an aspect or phenomenon, can lead to misconceptions or incomplete ideas about a concept or process.

In this dissertation we focus on how students' meta-modelling knowledge can be supported in secondary education. We specifically focus on a type of model that is often used in biology education: concept-process models. This kind of model displays biological processes, such as photosynthesis or cell division, and contains dynamics such as time and movement. Because of these dynamics, concept-process models cannot only be used to describe and simplify the real-world situation, but also to formulate hypotheses and predict future events. Consequently, these models can be seen as a suitable option for teachers to use when stimulating students' meta-modelling knowledge. The guiding research question for this dissertation was:

> *How can secondary school students' meta-modelling knowledge considering biological concept-process models be supported?*

In order to answer this question, we need both an instrument to measure students' meta-modelling knowledge and teaching and learning activities to support this knowledge development. Four studies are described in this dissertation that together discuss the development, application and evaluation of such an instrument and teaching and learning activities.

In **Chapter 2** we discuss the applicability of an existing framework containing important aspects of meta-modelling knowledge to biological concept-process models. The goal of this study was to find out whether this framework could be used to assess students' meta-modelling knowledge for this specific type of models. Dutch eleventh grade students from four different schools in the Netherlands were interviewed (n = 40, with 20 students on pre-university level and 20 students on higher general education level) to evaluate the applicability of the framework. This chapter was guided by the following research question:

> *2.1: To what extent can the described framework be used to assess students' understanding of biological concept-process models?*

Results showed that four additional descriptions were necessary to describe students' reasoning with biological concept-process models. The resulting framework shows five important aspects of meta-modelling knowledge (*nature of models, multiple models, purpose of models, testing models* and *changing models*) and four different levels of understanding, ranging from an initial level of understanding to an expert level of understanding.

To get a reliable picture of the current state of Dutch secondary students' expressed meta-modelling knowledge considering biological concept-process models, we used the framework from Chapter 2 to develop a test that could easily be distributed among many students throughout the Netherlands. The design of this test and its application are described in Chapter 3. Since the interviews that we carried out for the study as described in Chapter 2 indicated a possible difference in students' meta-modelling knowledge when focusing on models on a macroscopic level or on a (sub)microscopic level, we incorporated both types of models into our design. Also, since literature on this subject suggested a possible effect of presenting students with an explicit modelling-purpose on their expressed meta-modelling knowledge, we decided to also incorporate different types of modelling-purposes into the design. The following research questions were addressed in this study:

3.1: How does students' expressed meta-modelling knowledge, related to the aspects nature of models and multiple models, depend on the presence of different kinds of explicit modelling-purposes?

3.2: How does students' expressed meta-modelling knowledge, related to the aspects nature of models and multiple models, differ between contexts showing macroscopic and (sub)microscopic concept-process models?

The online test was filled out by 387 Dutch eleventh grade pre-university level students. Results showed that providing students with an explicit modelling purpose had no significant effect on their expressed level of meta-modelling knowledge considering biological concept-process models (RQ 3.1). Results for RQ 3.2 were in line with the results we found in the study described in Chapter 2, showing that students more often expressed a higher level of meta-modelling knowledge when presented with microscopic models than when presented with macroscopic models. However, the test showed that for both models students most often did not reach the highest levels of reasoning.

Chapter 4 discusses the development of the first lesson in a series of three lessons to support students' meta-modelling knowledge considering biological concept-process models. Lesson Study (LS) was used as a research approach for the development of this first lesson. LS originally is an approach where a team of teachers together develops and evaluates a lesson. In the study described in Chapter 4 we adapted the LS approach by adding researchers to the team of teachers that designs the lesson. The LS-team for this study consisted of three teachers, two researchers and a teacher who was also a researcher. In this study we developed teaching and learning activities that combined educational theory on important aspects of model-based reasoning with the pedagogical and didactical knowledge from teachers. In this first lesson students were introduced to important aspects of meta-modelling knowledge as described in the framework (Chapter 2). The following two research questions were addressed in this study:

4.1: To what extend does the developed lesson successfully familiarize students with important levels and aspects that are associated with model-based reasoning?

4.2: How do teachers and researchers experience using LS as a research approach?

During the developed lesson eleventh grade pre-university students (n = 34) were confronted with four models of the same biological process. They named differences between these models (Key activity 1, relating to the aspect *multiple models*), matched aims of a model to the different models (Key activity 2, relating to the aspect *purpose of models*), and formulated choices that the creator of the model had made in order to meet the aim of the model (Key activity 3, relating to the aspect *nature of models*). The developed lesson successfully familiarized students with the three important aspects of meta-modelling knowledge that were introduced to the students (*nature of models, multiple models* and *purpose of models*) (RQ 4.1). Considering RQ 4.2, three main focal points could be formulated for using LS as a research approach: 1) make sure that the teachers support the research question that the researchers bring into the LS-cycle, 2) take into account that the lesson is supposed to answer a research question which might cause extra stress for the teachers in a LS-team and 3) state the role of both researchers and teachers in a LS-team clearly at the beginning of the design process.

The focal points from Chapter 4 were taken into account when Lesson Study was used again as a research approach to develop two more lessons to stimulate students' reasoning with biological concept-process models. **Chapter 5** describes the design process and execution of these two lessons, in which the first lesson addresses the skill to read and write visual or symbolic language (i.e., *visual literacy*) and the second lesson focusses on model-based reasoning. The lessons were designed by the same LS-team as the one described in Chapter 4 and were also carried out in the same biology classes with the same students. The design was based on suggested interventions as described by Quillin and Thomas (2015), and the research was guided by the following two research questions:

> *5.1: What key activities are developed by the LS-team to foster students' visual literacy and engage students in modelling processes with biological concept-process models when following the suggested interventions by Quillin and Thomas (2015)?*
>
> *5.2: What is the influence of the developed key activities on students' reasoning with biological concept-process models?*

Both of the developed lessons consisted of multiple key activities, e.g., giving meaning to colours and arrows and creating a model of a biological process (RQ 5.1). Results from pre- and post-tests and from student interviews showed that the lessons on visual literacy and model-based reasoning contributed to a more scientific view on the use of concept-process models in science considering the meta-modelling aspects *nature of models* and *multiple models* (RQ 5.2).

In our final study, described in **Chapter 6**, we extend our research to three-dimensional (3D) biological concept-process models. In this study we focus on whether students are able to apply the knowledge that they gained while working with two-dimensional (2D) concept-process models to a combination of 2D models and a 3D virtual reality (VR) model. For this study we developed a VR application on the topic of blood glucose regulation. After working with this application, students used a combination of 2D models and the VR model to answer questions related to model-based reasoning. One group of twelfth grade pre-university students worked with important aspects of meta-modelling knowledge before, in the form of the developed lessons as described in Chapter 4 and Chapter 5 (n = 41). Another group of twelfth grade pre-university students also worked with the application and filled out the test, but did not receive this form of prior instruction (n = 47). The following research question was addressed in this study:

> *6.1: Does prior instruction in meta-modelling result in differences in applying meta-modelling knowledge to a combination of 2D models and a 3D virtual environment*?

Results showed that students' knowledge about important aspects of meta-modelling knowledge related to 2D models also leads to higher levels of expressed meta-modelling knowledge when working with a combination of 2D models and the VR model. This suggests that students are able to extend the meta-modelling knowledge that they gained while working with 2D models to a 3D VR model.

All in all, this dissertation sheds light on a possible way to stimulate students' meta-modelling knowledge. The adapted version of LS, where we added researchers to the LS-team to develop lessons on model-based reasoning, has shown to be a promising option for bridging the gap between theory-driven research and educational practice in this area. Using the test we developed that specifically focuses on biological concept-process models, we found that the developed lessons resulted in students expressing a higher level of meta-modelling knowledge for the aspects *nature of model* and *multiple models*. Our research also brings up some focal points for incorporating model-based practices into the classroom. First, to support students in using language that is associated with model-based reasoning we recommend to not only follow the suggested interventions by Quillin and Thomas (2015) to foster *visual literacy* and model-based reasoning, but also to familiarize students with the five aspects of meta-modelling knowledge and the language that is associated with these aspects. Examples of key activities to familiarize students with the aspects *multiple models, purpose of models* and *nature of models* are naming differences between various models of the same process (*multiple models*), matching possible aims of a model to different models of the same process (*purpose of models*), and formulating choices that the creator of a model has made in order to meet the aim of the model (*nature of models*). Second, teachers need to gain knowledge and experience on model-based reasoning before they can teach their students about this topic. Third, it is important to implement model-based practices into the curriculum on a regular basis. Model-based reasoning is a meta-cognitive skill, and in order to master such skills, practice on a regular basis is necessary.

Future research is necessary to further validate our results. We especially encourage further research considering the use of LS as a research approach, in order to find out how sustainable the incorporation of theory into educational practice is when using this approach. Research can focus on the long-term effect of having been part of the LS-team on teachers, the long-term effect of the developed lessons on students, the effect of the developed lessons on students and teachers when taught by a teacher who was not part of the LS-team, and the effect of students' gained meta-modelling knowledge on their expressed meta-modelling knowledge related to models used in other STEM courses. Also, to get a more complete picture of students' expressed meta-modelling knowledge, we suggest to extend the test we developed to include all five aspects of meta-modelling knowledge.

SAMENVATTING

Samenvatting in het Nederlands

Summary in Dutch

Het kunnen begrijpen van wetenschappelijke informatie is niet alleen belangrijk voor wetenschappers, maar ook voor niet-wetenschappers. Modellen met gegevens over recente wetenschappelijke ontdekkingen zijn te zien in het nieuws, in kranten en op sociale media. Een voorbeeld van hoe deze overdracht van informatie naar de samenleving plaatsvindt, is te zien bij de uitbraak van de COVID-19 pandemie in 2020. Overal duiken modellen op die laten zien hoe groepsimmuniteit mensen kan beschermen, hoe vaccinatie werkt, of wat de prognose is met betrekking tot het verloop van de pandemie. Deze verspreiding van wetenschappelijke kennis in de vorm van modellen naar de samenleving maakt het voor iedereen belangrijk om het ontstaan en het gebruik van modellen in de wetenschap te begrijpen.

Het onderwijs kan een belangrijke rol spelen bij het aanleren van kennis over modellen; zowel wat betreft het aanleren van kennis over de verschillende manieren waarop een model gebruikt kan worden, als het aanleren van kennis over het wetenschappelijk proces dat plaatsvindt bij het maken van een model. Dergelijke kennis wordt ook wel *meta-modelling knowledge* genoemd. Om het verwerven van deze kennis te stimuleren, hebben veel landen het zelf maken van modellen als activiteit in hun curriculum opgenomen. Uit de literatuur over dit onderwerp blijkt echter dat docenten in de praktijk vooral modellen gebruiken als hulpmiddel bij het illustreren van theorie, niet om de mogelijkheden van het gebruik van modellen of de wetenschappelijke praktijken achter de modellen te belichten. Leerlingen wordt daarom vaak niet expliciet de kennis over modellen bijgebracht die nodig is om het gebruik van modellen in de wetenschap te begrijpen. Leerlingen leren bijvoorbeeld niet dat modellen vaak vereenvoudigingen van de werkelijkheid zijn, en alleen weergeven wat op dat moment bekend is of verondersteld wordt over een bepaald aspect of fenomeen. Het ontbreken van dit soort kennis kan tot misvattingen of onvolledige ideeën over een concept of proces leiden.

In dit proefschrift richten we ons op hoe de ontwikkeling van *meta-modelling knowledge* van leerlingen kan worden ondersteund in het voortgezet onderwijs. We richten ons specifiek op een type model dat veel gebruikt wordt in het biologieonderwijs: het concept-procesmodel. Dit type modellen geeft biologische processen weer, zoals fotosynthese of celdeling. Het is een dynamisch model, waarin concepten en processen in de tijd worden weergegeven. Vanwege dit dynamische aspect, kunnen concept-procesmodellen niet alleen worden gebruikt om de werkelijke situatie te beschrijven en te vereenvoudigen, maar ook om hypotheses te formuleren en toekomstige gebeurtenissen te voorspellen. Hierdoor kunnen deze modellen gezien worden als een geschikte optie voor docenten om de *meta-modelling knowledge* van leerlingen de stimuleren. De leidende onderzoeksvraag voor dit proefschrift luidt:

> *Hoe kan de meta-modelling knowledge van leerlingen in het voortgezet onderwijs met betrekking tot biologische concept-proces modellen worden ondersteund?*

Om deze vraag te beantwoorden hebben we zowel een instrument nodig om de *meta-modelling knowledge* van leerlingen zichtbaar te maken, als onderwijsleeractiviteiten om deze kennisontwikkeling te ondersteunen. De vier studies in dit proefschrift beschrijven de ontwikkeling, toepassing en evaluatie van een instrument om de *meta-modelling knowledge* van leerlingen zichtbaar te maken, en de ontwikkeling van onderwijsleeractiviteiten om *meta-modelling* knowledge bij leerlingen te ondersteunen.

In **Hoofdstuk 2** bespreken we of een bestaand raamwerk dat belangrijke aspecten van *meta-modelling knowledge* beschrijft toepasbaar is op biologische concept-procesmodellen. Het doel van deze studie was om na te gaan of dit raamwerk gebruikt kan worden om de *meta-modelling knowledge* van leerlingen voor dit specifieke type modellen te beoordelen. Nederlandse leerlingen uit de vijfde klas van vier verschillende scholen in Nederland werden geïnterviewd (n = 40, waarvan 20 leerlingen op vwo-niveau en 20 leerlingen op havo-niveau) om de toepasbaarheid van het raamwerk te evalueren. Deze studie werd geleid door de volgende onderzoeksvraag:

> *2.1: In hoeverre kan het beschreven raamwerk gebruikt worden om het begrip van leerlingen van biologische concept-proces modellen te beoordelen?*

Uit de resultaten bleek dat er vier aanvullende beschrijvingen nodig waren om het redeneren van leerlingen met biologische concept-procesmodellen te beschrijven. Het resulterende raamwerk toont vijf belangrijke aspecten van *meta-modelling knowledge* (aard van modellen, meerdere modellen, doel van modellen, testen van modellen en veranderen van modellen) en vier verschillende niveaus van begrip, variërend van een beginnend begripsniveau tot een expert begripsniveau. Als voorbeeld: bij het aspect 'aard van modellen' zit een leerling op een beginnend begripsniveau als hij denkt dat modellen exacte kopieën van de werkelijkheid zijn, en op een expert begripsniveau als hij ziet dat modellen een hypothetische weergave van (een deel van) de werkelijkheid zijn.

In **Hoofdstuk 3** beschrijven we de studie die we uitvoerden om een betrouwbaar beeld van de huidige *meta-modelling knowledge* van Nederlandse middelbare scholieren te krijgen, waarbij we focusten op biologische concept-proces modellen. Voor deze studie hebben we het raamwerk uit Hoofdstuk 2 gebruikt om een digitale vragenlijst te ontwikkelen die gemakkelijk verspreid kon worden onder een groot aantal scholieren in heel Nederland. Zowel het ontwerp van deze vragenlijst als de toepassing ervan worden beschreven in Hoofdstuk 3. Omdat de resultaten van de studie zoals beschreven in Hoofdstuk 2 ons deden vermoeden dat er een mogelijk verschil in *meta-modelling knowledge* bij leerlingen was wanneer zij zich richten op modellen op macroscopisch niveau of op (sub)microscopisch niveau, hebben we beide typen modellen in ons ontwerp opgenomen. Ook hebben we verschillende doelen van modellen in het ontwerp van de vragenlijst opgenomen, aangezien de literatuur over dit onderwerp een mogelijk effect suggereerde van het presenteren van een expliciet doel van het model op de *meta-modelling knowledge* die leerlingen laten zien. De volgende onderzoeksvragen werden in deze studie behandeld:

> *3.1: Hoe hangt de geuite meta-modelling knowledge van leerlingen, gerelateerd aan de aspecten aard van modellen en meerdere modellen, af van de aanwezigheid van verschillende soorten expliciete model doeleinden?*
>
> *3.2: Welk verschil in geuite meta-modelling knowledge van leerlingen, gerelateerd aan de aspecten aard van modellen en meerdere modellen, is er tussen contexten met macroscopische en (sub)microscopische concept-procesmodellen?*

De online vragenlijst werd ingevuld door 387 Nederlandse vwo-leerlingen uit de vijfde klas. De resultaten toonden aan dat het expliciet meegeven van een doel van een model aan studenten geen significant effect had op het geuite niveau van *meta-modelling knowledge* met betrekking tot biologische concept-procesmodellen (de eerste onderzoeksvraag). De resultaten voor de tweede onderzoeksvraag waren in overeenstemming met de resultaten die we vonden in de studie beschreven in Hoofdstuk 2, waaruit bleek dat leerlingen vaker een hoger niveau van *meta-modelling knowledge* lieten zien wanneer ze (sub) microscopische modellen voorgelegd kregen dan wanneer ze macroscopische modellen voorgelegd kregen. Uit de vragenlijst bleek echter dat leerlingen voor beide type modellen meestal niet het hoogste niveau van redeneren lieten zien.

Hoofdstuk 4 bespreekt de ontwikkeling van een les ter ondersteuning van de *meta-modelling knowledge* van leerlingen met betrekking tot biologische concept-procesmodellen. Lesson Study (LS) werd gebruikt als onderzoeksmethode voor de ontwikkeling van deze eerste les. LS is van oorsprong een methode waarbij een team van docenten samen een les ontwikkelt en evalueert. In de studie beschreven in Hoofdstuk 4 hebben we de LS-benadering aangepast door onderzoekers toe te voegen aan het team van docenten dat de les ontwerpt. Het LS-team voor deze studie bestond uit drie biologie docenten van één school, twee onderzoekers en een biologie docent die tevens onderzoeker was. In deze studie ontwikkelden we onderwijsleeractiviteiten die de onderwijstheorie over belangrijke aspecten van modelmatig redeneren combineerden met de pedagogische en didactische kennis van docenten. In deze eerste les maakten leerlingen kennis met belangrijke aspecten van *meta-modelling knowledge* zoals beschreven in het raamwerk (Hoofdstuk 2). De volgende twee onderzoeksvragen werden in deze studie beantwoord:

> *4.1: In hoeverre maakt de ontwikkelde les leerlingen succesvol bekend met belangrijke niveaus en aspecten die samenhangen met meta-modelling knowledge?*
>
> *4.2: Hoe ervaren docenten en onderzoekers het gebruik van Lesson Study als onderzoeksmethode?*

Om de eerste vraag te beantwoorden kregen vijfdejaars vwo-leerlingen (n = 34) tijdens de ontwikkelde les vier modellen van hetzelfde biologische proces voorgelegd. Ze benoemden verschillen tussen deze modellen (kernactiviteit 1, met betrekking tot het aspect 'meerdere modellen'), koppelden doelen van een model aan de verschillende modellen (kernactiviteit 2, met betrekking tot het aspect 'doel van modellen'), en formuleerden keuzes die de maker van het model had gemaakt om aan het doel van het model te voldoen (kernactiviteit 3, met betrekking tot het aspect 'aard van modellen'). De resultaten lieten zien dat de ontwikkelde les de leerlingen met succes bekend gemaakt had met de drie belangrijke aspecten van *meta-modelling knowledge* die in de les aan bod kwamen: aard van modellen, meerdere modellen en doel van modellen.

Wat betreft de tweede onderzoeksvraag konden drie belangrijke aandachtspunten worden geformuleerd voor het gebruik van LS als onderzoeksaanpak: 1) zorg ervoor dat de docenten de onderzoeksvraag ondersteunen die de onderzoekers inbrengen in de LS-cyclus, 2) houd er rekening mee dat het feit dat met de les een onderzoeksvraag beantwoord dient te worden extra stress kan opleveren voor de docenten in een LS-team en 3) bespreek de rol van zowel onderzoekers als docenten in een LS-team duidelijk bij het begin van het ontwerpproces.

De aandachtspunten uit Hoofdstuk 4 zijn meegenomen in de studie die in **Hoofdstuk 5** besproken wordt, waarin Lesson Study opnieuw als onderzoeksmethode wordt ingezet om twee lessen te ontwikkelen om het redeneren van leerlingen met biologische concept-procesmodellen verder te stimuleren. Hoofdstuk 5 beschrijft zowel het ontwerpproces als de uitvoering van deze twee lessen, waarbij de eerste les ingaat op de vaardigheid om visuele of symbolische taal te lezen en te schrijven (i.e., visuele geletterdheid) en de tweede les zich richt op het redeneren met modellen. De lessen werden ontworpen door hetzelfde LS-team als het team dat beschreven is in Hoofdstuk 4. De lessen werden ook uitgevoerd in dezelfde biologieklassen, met dezelfde leerlingen, en in hetzelfde schooljaar. Het ontwerp was gebaseerd op voorgestelde interventies zoals beschreven door Quillin en Thomas (2015) en het onderzoek werd geleid door de volgende twee onderzoeksvragen:

> *5.1: Welke kernactiviteiten worden door het Lesson Study team ontwikkeld om de visuele geletterdheid van leerlingen te bevorderen en ze te betrekken bij het redeneren met biologische concept-procesmodellen wanneer de voorgestelde interventies van Quillin en Thomas (2015) worden gevolgd?*
>
> *5.2: Wat is de invloed van de ontwikkelde kernactiviteiten op de meta-modelling knowledge van leerlingen?*

Beide ontwikkelde lessen bestonden uit meerdere kernactiviteiten, bv. betekenis geven aan kleuren en pijlen en een model maken van een biologisch proces (eerste onderzoeksvraag). Resultaten van pre- en post-tests en van de interviews met leerlingen toonden aan dat de lessen over visuele geletterdheid en modelmatig redeneren bijdroegen aan een meer wetenschappelijke kijk op het gebruik van concept-procesmodellen in de wetenschap wanneer gelet wordt op de aspecten aard van modellen en meerdere modellen (tweede onderzoeksvraag).

In onze laatste studie, beschreven in **Hoofdstuk 6**, breidden we ons onderzoek uit naar driedimensionale (3D) biologische concept-procesmodellen. In dit onderzoek richtten we ons op de vraag of studenten in staat zijn om de kennis die ze hebben opgedaan tijdens het werken met tweedimensionale (2D) concept-procesmodellen toe te passen op een combinatie van 2D modellen en een 3D virtual reality (VR) model. Voor dit onderzoek ontwikkelden we een VR applicatie over het onderwerp bloedglucoseregulatie. Na het werken met deze applicatie gebruikten leerlingen een combinatie van 2D modellen en het VR model om vragen te beantwoorden met betrekking tot het redeneren met modellen. Met behulp van een pre- en post-test werd gekeken wat het effect van deze interventie was op de *meta-modelling knowledge* van leerlingen. Eén groep zesdejaars vwo-leerlingen werkte al eerder met belangrijke aspecten van *meta-modelling knowledge*, in de vorm van de ontwikkelde lessen zoals beschreven in Hoofdstuk 4 en Hoofdstuk 5 (n = 41). Een andere groep vwo-leerlingen uit de zesde klas fungeerde als controlegroep. Deze groep leerlingen heeft ook met de applicatie gewerkt en de vragenlijst ingevuld, maar heeft de eerdere instructie in de vorm van de ontwikkelde lessen niet gekregen (n = 47). De volgende onderzoeksvraag werd voor deze studie geformuleerd:

> *6.1: Wat is het effect van voorafgaande instructie op het gebied van meta-modelling knowledge op het toepassen van dergelijke meta-modelling knowledge door leerlingen op een combinatie van 2D modellen en een 3D virtuele omgeving?*

De resultaten toonden aan dat de kennis van leerlingen over belangrijke aspecten van *meta-modelling knowledge* met betrekking tot 2D modellen ook leidt tot hogere niveaus van geuite *meta-modelling knowledge* bij het werken met een combinatie van 2D modellen en het VR model. Dit suggereert dat leerlingen in staat zijn om de *meta-modelling knowledge* die ze hebben opgedaan bij het werken met 2D modellen uit te breiden naar een 3D VR model.

Samenvattend laat dit proefschrift een manier zien om de *meta-modelling knowledge* van leerlingen te stimuleren. De aangepaste versie van LS, waarbij we onderzoekers aan het LS-team hebben toegevoegd om lessen over *meta-modelling knowledge* te ontwikkelen, blijkt een veelbelovende optie om de kloof tussen theorie-gedreven onderzoek en de onderwijspraktijk op dit gebied te overbruggen. Daarnaast hebben we met behulp van de door ons ontwikkelde vragenlijst die specifiek gericht is op biologische concept-procesmodellen vastgesteld dat de ontwikkelde lessen ertoe leidden dat leerlingen een hoger niveau van *meta-modelling knowledge* tot uitdrukking brachten voor de aspecten aard van modellen en meerdere modellen.

Ons onderzoek brengt ook enkele aandachtspunten naar voren voor het implementeren van activiteiten die gerelateerd zijn aan het redeneren met modellen in de klas. Ten eerste, om leerlingen te ondersteunen in het gebruik van taal die geassocieerd wordt met modelgebaseerd redeneren, raden we aan om niet alleen de voorgestelde interventies van Quillin en Thomas te volgen om visuele geletterdheid en het redeneren met modellen te bevorderen, maar ook om leerlingen vertrouwd te maken met de vijf aspecten van *meta-modelling knowledge* en de taal die geassocieerd wordt met deze aspecten zoals beschreven in het raamwerk in Hoofdstuk 2. Voorbeelden van kernactiviteiten om leerlingen vertrouwd te maken met de aspecten meerdere modellen, doel van modellen en aard van modellen zijn het benoemen van verschillen tussen verschillende modellen van hetzelfde proces (meerdere modellen), het afstemmen van mogelijke doelen van een model op verschillende modellen van hetzelfde proces (doel van modellen), en het formuleren van keuzes die de maker van een model heeft gemaakt om aan het doel van het model te voldoen (aard van modellen). Ten tweede is het belangrijk dat leraren zelf kennis en ervaring opdoen met het redeneren met modellen en de aspecten van *meta-modelling knowledge* voordat zij hun leerlingen hierover kunnen onderwijzen. Ten derde is het belangrijk om het redeneren met modellen op regelmatige basis in het curriculum terug te laten komen. Redeneren met modellen is een meta-cognitieve vaardigheid en om een dergelijke vaardigheid onder de knie te krijgen is regelmatige oefening noodzakelijk.

Toekomstig onderzoek is nodig om onze resultaten verder te valideren. Wij moedigen in het bijzonder verder onderzoek aan naar het gebruik van Lesson Study als onderzoeksmethode, om uit te vinden hoe duurzaam de implementatie van theorie in de onderwijspraktijk is wanneer LS op deze manier wordt ingezet. Onderzoek kan zich richten op het langetermijneffect van het deel uitmaken van een dergelijk LS-team op docenten, het langetermijneffect van de ontwikkelde lessen op leerlingen, het effect van de ontwikkelde lessen op leerlingen en docenten wanneer zij les krijgen van een docent die geen deel uitmaakte van het LS-team, en het effect van de opgedane *meta-modelling knowledge* van leerlingen op hun geuite *meta-modelling knowledge* met betrekking tot modellen die in andere bèta/technische vakken worden gebruikt. Om een vollediger beeld te krijgen van de meta-modelling knowledge van leerlingen, stellen wij daarnaast voor om de door ons ontwikkelde vragenlijst uit te breiden met alle vijf aspecten van *meta-modelling knowledge*.

DANKWOORD

Dankwoord

Tijdens het schrijven van mijn masterscriptie begonnen mensen wel eens over het fenomeen promoveren. ‘Ik ga echt nooit beginnen aan een PhD’, bleef ik stug volhouden. Zo’n scriptie schrijven was al gedoe, dan moet zo’n proefschrift schrijven echt verschrikkelijk zijn, dacht ik. Dat niet-beginnen aan een PhD heb ik een aantal jaar volgehouden. Ik gaf biologieles op het Gymnasium Novum in Voorburg, en werkte daarnaast bij de Praktijk in Amsterdam, een creatief bureau voor wetenschapscommunicatie en onderwijs. Alex Verkade was destijds directeur bij de Praktijk en gaf aan dat een vriend van hem op zoek was naar een promovendus. Het leek hem echt iets voor mij. Enigszins sceptisch, maar ook zeker vereerd, maakte ik een afspraak om te horen waar het project over zou gaan. Zo ontmoette ik bij café Oudaen Wouter van Joolingen, die mij onder het genot van een biertje uitlegde wat voor project hij voor ogen had. Het was een gezellig gesprek, waarbij we binnen no-time enthousiast met behulp van bierviltjes en glazen onze ideeën aan elkaar aan het uitbeelden waren. Die middag ging ik naar huis met het gevoel dat de tijd waarin ik niet ging beginnen aan een PhD wellicht tot een einde was gekomen.

Na het schrijven van het projectvoorstel en het binnenhalen van een NWO lerarenbeurs was het in 2016 dan echt zo ver, mijn promotietraject ging van start. Het was een rollercoaster. Het combineren van drie dagen per week op school en twee dagen per week op de universiteit ging niet zonder slag of stoot. Vergaderingen op universiteitsdagen, dataverzamelmomenten op schooldagen, alles liep door elkaar heen. Maar het was het meer dan waard, want wat heb ik veel geleerd! Kwalitatief onderzoek doen, het presenteren van je werk op congressen, en het omgaan met andere culturen tijdens de summerschool en op congressen in verschillende landen, zijn enkele voorbeelden van voor mij zeer leerzame momenten. Daarnaast heeft dit promotietraject me heel veel gezellige momenten en nieuwe mensen in mijn leven opgeleverd. De reis naar Singapore waar ik samen met het VR lab op Nanyang Technological University (NTU) een app ontwikkelde zie ik als een enorme plus. Het werk heeft een mooie publicatie opgeleverd, maar voor mij persoonlijk is ook het verblijf van een aantal maanden in een andere cultuur zeer waardevol geweest. Het heeft me (met vallen en opstaan) veel geleerd over de Singaporese bevolking, en heeft geleid tot vriendschappen die zelfs na thuiskomst in Nederland stand hebben gehouden. Zonder de hulp en steun van heel veel verschillende mensen had ik dit promotietraject nooit tot een goed einde kunnen brengen. Deze mensen wil ik dan ook graag bedanken.

Allereerst Wouter, mijn promotor. Brainstormen met jou over mijn onderzoek voelde als iets natuurlijks. Na een gesprek met jou wist ik altijd dat het inhoudelijk weer een stukje beter was geworden. Daarnaast was jij vaak degene die als een taalvirtuoos de *smashing eindzin* bij artikelen componeerde. Ook denk ik met veel plezier terug aan de etentjes en de goede gesprekken die we hadden over alle belangrijke zaken in het leven.

Christine, mijn copromotor en beschermengel. Je was ontzettend zorgvuldig in het bekijken van de door mij geschreven teksten en kon steeds weer precies de vinger op de zere plek leggen. Naast het inhoudelijke werk, hield je altijd scherp in de gaten of het wel goed met me ging en ik niet weer eens te veel hooi op mijn vork had genomen. Bedankt voor alle fijne gesprekken en het altijd oplettende oog.

Michiel en Melde, mijn paranimfen, mede-PhD-ers en biohotties. Michiel, van onderwijskundige visie tot favoriete hagelslag, over alles lijken we hetzelfde te denken. Samen met jou heb ik met veel plezier de PhD-weekenden gepland, presentaties gegeven, wandelingen door de botanische tuin gemaakt en gedanst op congressen. Melde, het grootste deel van mijn promotietraject zat ik tegenover jou op een kamer op het FI. We deelden lief en leed, konden inhoudelijk fijn met elkaar sparren, en hebben voorafgaand aan de summerschool heerlijk een weekend samen Helsinki verkend. Ik heb ontzettend veel aan onze gesprekken gehad. Michiel en Melde, ik vind het heel fijn dat jullie mijn paranimfen zijn en we ook tijdens mijn verdediging een team vormen.

I would like to thank Prof. Cai and his team at NTU for developing the VR app that we used in our final study. Especially Ryan and Siti who spent so much time explaining to me what can and cannot be done in VR, while incorporating my ideas in the app. I would also like to thank Siti and Zhe An for showing me around and making me feel at home in Singapore.

Ook de andere collega's en mede PhD-ers op het Freudenthal Instituut wil ik graag bedanken voor hun betrokkenheid. Met name de gesprekken met Winnifred die me met haar scherpzinnige vragen altijd weer een spiegel wist voor te houden, and Luhuan who always showed interest in Dutch culture and brought back the sweetest gifts from China zijn mij veel waard geweest.

De schoolleiding van het Gymnasium Novum wil ik bedanken dat zij mij de gelegenheid hebben geboden om in deeltijd te promoveren. Mijn onderzoek heeft zich gedeeltelijk ook op deze school afgespeeld. Samen met Willemien, Harm en Marieke vormden we een ontzettend sterk Lesson Study team. Ik wil jullie drieën dan ook bedanken voor alle tijd die jullie in dit onderzoek hebben gestoken, voor jullie creatieve ideeën, en voor het samen vormgeven van lessen. Harm, jou wil ik met name in het zonnetje zetten vanwege het feit dat jij alle ontworpen lessen onder toeziend oog van het hele team durfde te geven. Met betrekking tot het uitvoeren van het onderzoek wil ik naast het Gymnasium Novum ook alle andere middelbare scholen en hun leerlingen bedanken die hun tijd beschikbaar hebben gesteld om een bijdrage te leveren aan mijn onderzoek.

Buiten werktijd zijn er natuurlijk ook veel mensen geweest die mij hebben gesteund, hebben opgevangen, of me er af en toe doorheen hebben gesleept. Martine en Silvia, mijn snowchicks. Er zijn geen vriendinnen die meer van mijn dagelijkse sores meekrijgen dan jullie. Met onbevooroordeelde houding zijn jullie er altijd om te luisteren naar zo ongeveer alles waar ik tegenaan loop in het leven. Wat betreft dit promotietraject had ik het zonder het met jullie kunnen delen van mijn paniekmomenten nooit gered.

Alle andere vrienden en familie, met name Julia, Nathalie, Sioe-Lie, Tineke, Yukari, Laura, Leoniek, Richard, Dominik en Aart, bedankt dat ik met jullie altijd de mooie én de minder mooie momenten in het leven kan delen. Het is een rijkdom om zulke fijne vrienden om mij heen te hebben.

Serge en Alexandra, mijn schoonouders, wat is het fijn om jullie met jullie inspirerende interesse en enthousiasme in mijn leven te hebben. Bedankt voor alle goede gesprekken en hulp thuis als het allemaal toch even een beetje te veel werd.

Papa en mama en mijn zusje Simone, bedankt voor het steeds blijven doorvragen en het altijd maar weer proberen te begrijpen wat ik nu eigenlijk precies aan het doen was. Jullie waren zowel het perfecte lekenpubliek, als het warme bad waar ik terecht kon als ik er even helemaal doorheen zat.

Sven, mijn liefste, jij kent mij door en door en weet altijd, ook wanneer ik niets zeg, wat er speelt. Je maakt me aan het lachen, relativeert (wanneer ik daar om vraag), en gaat samen met mij graag het avontuur aan. Ook dit promotietraject was een avontuur, een periode waarin je mij van zowel mijn beste als mijn slechtste kant hebt gezien. Bedankt voor je steun en voor het feit dat jij altijd in mij gelooft. Tijdens mijn promotietraject is onze lieve dochter Eva geboren en dat heeft van ons een gezinnetje gemaakt. Jou als papa zien vervult me met liefde en trots. Ik kijk uit naar alle avonturen die wij nog met zijn drietjes gaan meemaken.

Curriculum Vitae

Susanne Jansen was born on March 24, 1986 in Schiedam, the Netherlands. She completed her secondary education at Scholengemeenschap Spieringshoek in Schiedam in 2004. In that same year she moved to Utrecht and started the bachelor Biology at Utrecht University. After incorporating a minor on cognitive neurosciences into this bachelor, she completed the bachelor in 2008. The elective courses on cellular and molecular biology during this bachelor got her interested in biomedical sciences. She therefore chose to combine her interests in molecular biology and neurosciences during an optional internship at MGH Center for Human Genetic Research in Boston in 2008. During this internship she focused on the influence of Single Nucleotide Polymorphisms (SNP's) on various psychological diseases. In 2009 she started the master Biomedical Sciences at the University of Amsterdam, following the track medical biology. She graduated in 2012 with a study on the use of oral contraceptives in BRCA1/2 mutation carriers. In the final years of her study, Susanne realised that teaching brought her more joy than working in a laboratory. She therefore started a second master at the University of Amsterdam to obtain a teaching degree, which she also finished in 2012.

Susanne started working at Gymnasium Novum as a biology teacher in 2011. In 2012 she combined this work with a job at de Praktijk, a creative agency for science communication and education, where she designed teaching material for students and developed workshops for teachers. In 2013 she decided to work fulltime at Gymnasium Novum, where she combined teaching with several projects that focused on differentiation in the classroom and exchanging knowledge between teachers. In 2016 she received a grant from the Dutch Research Council (NWO, project number: 023.007.065). This enabled the start of her PhD project at Utrecht University on fostering students' meta-modelling knowledge regarding biological concept-process models. In addition to conducting her research, Susanne continued her work at Gymnasium Novum. She also started working at de Vierde Opleidingsschool, a collaboration of several teacher training institutes and secondary schools with the aim to educate and coach students on their journey to become teachers.

Susanne lives in Utrecht with her partner Sven and their daughter Eva. In her free time she enjoys reading a good book, bouldering and playing the piano.

Publications related to this dissertation

Jansen, S., Knippels, M.C.P.J., & van Joolingen, W.R. (2021), Lesson Study as a research approach: a case study, *International Journal for Lesson and Learning Studies, 10*(3), 286 – 301. https://doi.org/10.1108/IJLLS-12- 2020-0098

Jansen, S., Knippels, M.C.P.J., & van Joolingen, W.R. (2019), Assessing students' understanding of models of biological processes: a revised framework, *International Journal of Science Education*, *41*(8), 981- 994. https://doi.org/10.1080/09500693.2019.1582821

Jansen, S., Aung, R., Binte Mohd Taib, S.F., Cai, Y., & van Joolingen, W.R. (accepted for publication). Applying prior meta-modelling knowledge to a VR model of a biological process. In Y. Cai., E. Mangina & S. L. Goei (Eds.), *Mixed Reality for Education*. Springer.

Jansen, S., Knippels, M.C.P.J. and van Joolingen, W.R., *Fostering students' meta-modeling knowledge about biological concept-process models*. [Manuscript submitted for publication]

Jansen, S., Knippels, M.C.P.J., van Montfort, M. and van Joolingen, W.R., Secondary school students' meta-modeling knowledge of biological concept-process models. [Manuscript submitted for publication]

Presentations related to this dissertation

Jansen, S., Aung, R., Binte Mohd Taib, S.F., Cai, Y., & van Joolingen, W.R. (2020). *Introducing reasoning with VR models in biology education*. Symposium at the Centre for Augmented and Virtual Reality, Nanyang Technological University, Singapore, December 11.

Jansen, S., Knippels, M.C.P.J., & van Joolingen, W.R (2019). *Fostering scientific reasoning with models of biological processes*. Paper presentation at the World Association of Lesson Studies (WALS) conference, Amsterdam, the Netherlands, September 3-6.

Jansen, S., Knippels, M.C.P.J., & van Joolingen, W.R. (2019). *Fostering scientific reasoning with models of biological processes*. Paper presentation at the 13th ESERA conference, Bologna, Italy, August 26-30.

Jansen, S., Aung, R., Binte Mohd Taib, S.F., Cai, Y., & van Joolingen, W.R. (2019). *Introducing reasoning with VR models in biology education*. Symposium at the Centre for Augmented and Virtual Reality, Nanyang Technological University, Singapore, August 1.

Jansen, S., Knippels, M.C.P.J., & van Joolingen, W.R. (2019). *Developing higher order thinking skills using Lesson Study*. Webinar, coördinated by the World Association of Lesson Studies (WALS), Amsterdam, the Netherlands, July 25.

Jansen, S., Knippels, M.C.P.J., & van Joolingen, W.R. (2019). *Het ontwikkelen van hogere orde denkvaardigheden met behulp van Lesson Study [Developing higher order thinking skills using Lesson Study]*. Symposium at the ORD conference, Heerlen, the Netherlands, June 26-28.

Jansen, S., Knippels, M.C.P.J., & van Joolingen, W.R. (2018). *Assessing and improving students' understanding of models of biological processes*. Symposium at Humboldt-Universität, Berlin, July 18.

Jansen, S., Knippels, M.C.P.J., & van Joolingen, W.R. (2018). *Assessing students' understanding of models of biological processes*. Set-up and first results of the research project presented at the 15th ESERA Summer School, Jyväskylä, Finland, June 25 to 1 July.

Jansen, S., Knippels, M.C.P.J., & van Joolingen, W.R. (2017). *Biology curriculum & research*. First results of the research project presented at the Centre for Augmented and Virtual Reality, Nanyang Technological University, Singapore, October 15.

FI Scientific Library

(Formerly published as CD-b Scientific Library)

114. Pieters, M.L.M. (2022). *Between written and enacted: Curriculum development as propagation of memes. An ecological-evolutionary perspective on fifty years of curriculum development for upper secondary physics education in the Netherlands*

113. Veldkamp, A. (2022). *No Escape! The rise of escape rooms in secondary science education.*

112. Kamphorst, F. (2021). *Introducing Special Relativity in Secondary Education.*

111. Leendert, A.-M. J. M. van (2021). *Improving Reading and Comprehending Mathematical Expressions in Braille*.

110. Gilissen, M. G. R. (2021). *Fostering Students' System Thinking in Secondary Biology Education*.

109. Dijke-Droogers, M.J.S. van (2021). *Introducing Statistical Inference: Design and Evaluation of a Learning Trajectory.*

108. Wijnker, W. (2021). *The Unseen Potential of Film for Learning. Film's Interest Raising Mechanisms Explained in Secondary Science and Mathematics Education.*

107. Groothuijsen, S. (2021). *Quality and impact of practice-oriented educational research*.

106. Wal, N.J. van der (2020). *Developing Techno-mathematical Literacies inhigher technical professional education*.

105. Tacoma, S. (2020). *Automated intelligent feedback in university statisticseducation*.

104. Zanten, M. van (2020). *Opportunities to learn offered by primary school mathematics textbooks in the Netherlands*

103. Walma, L. (2020). *Between Morpheus and Mary: The Public Debate onMorphine in Dutch Newspapers, 1880-1939*

102. Van der Gronde, A.G.M.P. (2019). *Systematic Review Methodology inBiomedical Evidence Generation.*

101. Klein, W. (2018). *New Drugs for the Dutch Republic. The Commodification of Fever Remedies in the Netherlands (c. 1650-1800).*

100. Flis, I. (2018). *Discipline Through Method - Recent history and philosophy of scientific psychology (1950-2018).*

99. Hoeneveld, F. (2018). Een vinger in de Amerikaanse pap. Fundamenteelfysisch en defensie onderzoek in Nederland tijdens de vroege Koude Oorlog.

98. Stubbé-Albers, H. (2018). Designing learning opportunities for the hardest to reach: Game-based mathematics learning for out-of-school children in Sudan.

97. Dijk, G. van (2018). *Het opleiden van taalbewuste docenten natuurkunde,scheikunde en techniek: Een ontwerpgericht onderzoek*.

96. Zhao, Xiaoyan (2018). *Classroom assessment in Chinese primary schoolmathematics education.*

95. Laan, S. van der (2017). *Een varken voor iedereen. De modernisering van de Nederlandse varkensfokkerij in de twintigste eeuw.*

94. Vis, C. (2017). Strengthening local curricular capacity in international development cooperation.

93. Benedictus, F. (2017). *Reichenbach: Probability & the A Priori. Has the Baby Been Thrown Out with the Bathwater?*

92. Ruiter, Peter de (2016). *Het Mijnwezen in Nederlands-Oost-Indië 1850- 1950.*

91. Roersch van der Hoogte, Arjo (2015). *Colonial Agro-Industrialism. Science, industry and the state in the Dutch Golden Alkaloid Age, 1850-1950.*

90. Veldhuis, M. (2015). *Improving classroom assessment in primary mathematics education.*

89. Jupri, Al (2015). *The use of applets to improve Indonesian studentperformance in algebra.*

88. Wijaya, A. (2015). *Context-based mathematics tasks in Indonesia:Toward better practice and achievement.*

87. Klerk, S. (2015). Galen reconsidered. Studying drug properties and thefoundations of medicine in the Dutch Republic ca. 1550-1700.

86. Krüger, J. (2014). *Actoren en factoren achter het wiskundecurriculum sinds 1600.*

85. Lijnse, P.L. (2014). *Omzien in verwondering. Een persoonlijke terugblik op 40 jaar werken in de natuurkundedidactiek.*

84. Weelie, D. van (2014). *Recontextualiseren van het concept biodiversiteit*.

83. Bakker, M. (2014). *Using mini-games for learning multiplication and division: a longitudinal effect study.*

82. Ngô Vũ Thu Hằng (2014). *Design of a social constructivism-basedcurriculum for primary science education in Confucian heritage culture*.

81. Sun, L. (2014). *From rhetoric to practice: enhancing environmental literacy of pupils in China*.

80. Mazereeuw, M. (2013). The functionality of biological knowledge in the workplace. Integrating school and workplace learning about reproduction.

79. Dierdorp, A. (2013). *Learning correlation and regression within authenticcontexts.*

78. Dolfing, R. (2013). Teachers' Professional Development in Context-basedChemistry Education. Strategies to Support Teachers in DevelopingDomain-specific Expertise.

77. Mil, M.H.W. van (2013). *Learning and teaching the molecular basis of life*.

76. Antwi, V. (2013). *Interactive teaching of mechanics in a Ghanaianuniversity context*.

75. Smit, J. (2013). *Scaffolding language in multilingual mathematics classrooms.*

74. Stolk, M. J. (2013). Empowering chemistry teachers for context-based education. Towards a framework for design and evaluation of a teacher professional development programme in curriculum innovations.

73. Agung, S. (2013). Facilitating professional development of Madrasah chemistry teachers. Analysis of its establishment in the decentralized educational system of Indonesia.

72. Wierdsma, M. (2012). *Recontextualising cellular respiration*.

71. Peltenburg, M. (2012). *Mathematical potential of special education students.*

70. Moolenbroek, A. van (2012). *Be aware of behaviour. Learning and teaching behavioural biology in secondary education.*

69. Prins, G. T., Vos, M. A. J., & Pilot, A. (2011). *Leerlingpercepties van onderzoek & ontwerpen in het technasium.*

68. Bokhove, Chr. (2011). *Use of ICT for acquiring, practicing and assessingalgebraic expertise*.

67. Boerwinkel, D. J., & Waarlo, A. J. (2011). *Genomics education for decision-making. Proceedings of the second invitational workshop on genomics education, 2-3 December 2010*.

66. Kolovou, A. (2011). *Mathematical problem solving in primary school*.

65. Meijer, M. R. (2011). Macro-meso-micro thinking with structure-propertyrelations for chemistry. An explorative design-based study.

64. Kortland, J., & Klaassen, C. J. W. M. (2010). *Designing theory-based teaching-learning sequences for science. Proceedings of the symposiumin honour of Piet Lijnse at the time of his retirement as professor of Physics Didactics at Utrecht University*.

63. Prins, G. T. (2010). *Teaching and learning of modelling in chemistry education. Authentic practices as contexts for learning*.

62. Boerwinkel, D. J., & Waarlo, A. J. (2010). *Rethinking science curricula inthe genomics era. Proceedings of an invitational workshop*.

61. Ormel, B. J. B. (2010). *Het natuurwetenschappelijk modelleren van dynamische systemen. Naar een didactiek voor het voortgezet onderwijs*.

60. Hammann, M., Waarlo, A. J., & Boersma, K. Th. (Eds.) (2010). *The nature of research in biological education: Old and new perspectives on theoretical and methodological issues – A selection of papers presented at the VIIth Conference of European Researchers in Didactics of Biology*.

59. Van Nes, F. (2009). *Young children's spatial structuring ability and emerging number sense*.

58. Engelbarts, M. (2009). Op weg naar een didactiek voor natuurkunde- experimenten op afstand. Ontwerp en evaluatie van een via internet uitvoerbaar experiment voor leerlingen uit het voortgezet onderwijs.

57. Buijs, K. (2008). *Leren vermenigvuldigen met meercijferige getallen*.

56. Westra, R. H. V. (2008). *Learning and teaching ecosystem behaviourin secondary education: Systems thinking and modelling in authentic practices*.

55. Hovinga, D. (2007). Ont-dekken en toe-dekken: Leren over de veelvormige relatie van mensen met natuur in NME-leertrajecten duurzame ontwikkeling.

54. Westra, A. S. (2006). A new approach to teaching and learning mechanics.

53. Van Berkel, B. (2005). *The structure of school chemistry: A quest for conditions for escape*.

52. Westbroek, H. B. (2005). *Characteristics of meaningful chemistry education: The case of water quality*.

51. Doorman, L. M. (2005). *Modelling motion: from trace graphs toinstantaneous change*.

50. Bakker, A. (2004). *Design research in statistics education: on symbolizingand computer tools*.

49. Verhoeff, R. P. (2003). *Towards systems thinking in cell biology education*.

48. Drijvers, P. (2003). Learning algebra in a computer algebra environment.Design research on the understanding of the concept of parameter.

47. Van den Boer, C. (2003). *Een zoektocht naar verklaringen voor achterblijvende prestaties van allochtone leerlingen in het wiskundeonderwijs*.

46. Boerwinkel, D. J. (2003). *Het vormfunctieperspectief als leerdoel van natuuronderwijs. Leren kijken door de ontwerpersbril*.

45. Keijzer, R. (2003). *Teaching formal mathematics in primary education. Fraction learning as mathematising process*.

44. Smits, Th. J. M. (2003). Werken aan kwaliteitsverbetering van leerlingonderzoek: Een studie naar de ontwikkeling en het resultaat vaneen scholing voor docenten.

43. Knippels, M. C. P. J. (2002). *Coping with the abstract and complex natureof genetics in biology education – The yo-yo learning and teaching strategy.*

42. Dressler, M. (2002). *Education in Israel on collaborative management ofshared water resources.*

41. Van Amerom, B.A. (2002). *Reinvention of early algebra: Developmentalresearch on the transition from arithmetic to algebra.*

40. Van Groenestijn, M. (2002). *A gateway to numeracy. A study of numeracyin adult basic education.*

39. Menne, J. J. M. (2001). Met sprongen vooruit: een productief oefenprogramma voor zwakke rekenaars in het getallengebied tot 100 –een onderwijsexperiment.

38. De Jong, O., Savelsbergh, E.R., & Alblas, A. (2001). *Teaching for scientific literacy: context, competency, and curriculum.*

37. Kortland, J. (2001). *A problem-posing approach to teaching decision making about the waste issue.*

36. Lijmbach, S., Broens, M., & Hovinga, D. (2000). *Duurzaamheid als leergebied; conceptuele analyse en educatieve uitwerking.*

35. Margadant-van Arcken, M., & Van den Berg, C. (2000). *Natuur inpluralistisch perspectief – Theoretisch kader en voorbeeldlesmateriaal voor het omgaan met een veelheid aan natuurbeelden.*

34. Janssen, F. J. J. M. (1999). *Ontwerpend leren in het biologieonderwijs. Uitgewerkt en beproefd voor immunologie in het voortgezet onderwijs*.

33. De Moor, E. W. A. (1999). *Van vormleer naar realistische meetkunde Een historisch-didactisch onderzoek van het meetkundeonderwijs aan kinderen van vier tot veertien jaar in Nederland gedurende de negentiende en twintigste eeuw.*

32. Van den Heuvel-Panhuizen, M., & Vermeer, H. J. (1999). *Verschillen tussenmeisjes en jongens bij het vak rekenen-wiskunde op de basisschool – Eindrapport MOOJ-onderzoek.*

31. Beeftink, C. (2000). Met het oog op integratie – Een studie over integratievan leerstof uit de natuurwetenschappelijke vakken in de tweede fase van het voortgezet onderwijs.

30. Vollebregt, M. J. (1998). *A problem posing approach to teaching an initialparticle model.*

29. Klein, A. S. (1998). Flexibilization of mental arithmetics strategies on a different knowledge base – The empty number line in a realistic versus gradual program design.

28. Genseberger, R. (1997). Interessegeoriënteerd natuur- en scheikundeonderwijs – Een studie naar onderwijsontwikkeling op de Open Schoolgemeenschap Bijlmer.

27. Kaper, W. H. (1997). *Thermodynamica leren onderwijzen*.

26. Gravemeijer, K. (1997). *The role of context and models in the development of mathematical strategies and procedures.*

25. Acampo, J. J. C. (1997). *Teaching electrochemical cells – A study onteachers' conceptions and teaching problems in secondary education*.
24. Reygel, P. C. F. (1997). *Het thema 'reproductie' in het schoolvak biologie*.
23. Roebertsen, H. (1996). *Integratie en toepassing van biologische kennis– Ontwikkeling en onderzoek van een curriculum rond het thema 'Lichaamsprocessen en Vergift'.*
22. Lijnse, P. L., & Wubbels, T. (1996). *Over natuurkundedidactiek, curriculumontwikkeling en lerarenopleiding*.
21. Buddingh', J. (1997). *Regulatie en homeostase als onderwijsthema: eenbiologie-didactisch onderzoek*.
20. Van Hoeve-Brouwer G. M. (1996). *Teaching structures in chemistry – Aneducational structure for chemical bonding*.
19. Van den Heuvel-Panhuizen, M. (1996). *Assessment and realisticmathematics education.*
18. Klaassen, C. W. J. M. (1995). *A problem-posing approach to teaching thetopic of radioactivity.*
17. De Jong, O., Van Roon, P. H., & De Vos, W. (1995). *Perspectives onresearch in chemical education.*
16. Van Keulen, H. (1995). *Making sense – Simulation-of-research in organicchemistry education.*
15. Doorman, L. M., Drijvers, P. & Kindt, M. (1994). *De grafische rekenmachine in het wiskundeonderwijs.*
14. Gravemeijer, K. (1994). *Realistic mathematics education*.
13. Lijnse, P. L. (Ed.) (1993). *European research in science education*.
12. Zuidema, J., & Van der Gaag, L. (1993). *De volgende opgave van de computer.*
11. Gravemeijer, K., Van den Heuvel-Panhuizen, M., Van Donselaar, G., Ruesink, N., Streefland, L., Vermeulen, W., Te Woerd, E., & Van der Ploeg, D. (1993). *Methoden in het reken-wiskundeonderwijs, een rijke context voor vergelijkend onderzoek.*
10. Van der Valk, A. E. (1992). *Ontwikkeling in Energieonderwijs*.
9. Streefland, L. (Ed.) (1991). *Realistic mathematics education in primary schools.*
8. Van Galen, F., Dolk, M., Feijs, E., & Jonker, V. (1991). *Interactieve videoin de nascholing reken-wiskunde.*
7. Elzenga, H. E. (1991). *Kwaliteit van kwantiteit*.
6. Lijnse, P. L., Licht, P., De Vos, W., & Waarlo, A. J. (Eds.) (1990). *Relatingmacroscopic phenomena to microscopic particles: a central problem in secondary science education.*
5. Van Driel, J. H. (1990). *Betrokken bij evenwicht*.
4. Vogelezang, M. J. (1990). *Een onverdeelbare eenheid*.
3. Wierstra, R. F. A. (1990). *Natuurkunde-onderwijs tussen leefwereld envakstructuur.*
2. Eijkelhof, H. M. C. (1990). *Radiation and risk in physics education*.
1. Lijnse, P. L., & De Vos, W. (Eds.) (1990). *Didactiek in perspectief.*