Lesson study as a research approach: a case study

Susanne Jansen, Marie-Christine P.J. Knippels and Wouter R. van Joolingen
Freudenthal Institute, Utrecht University, Utrecht, The Netherlands

Received 13 January 2021
Revised 1 April 2021
2 April 2021
Accepted 3 April 2021

Abstract
Purpose – The purpose of this paper 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.
Design/methodology/approach – 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 the 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 LS cycle, (2) take into account that the lesson is supposed to answer a research question that might cause extra stress for the teachers in an LS team and (3) state the role of both researchers and teachers in an LS team clearly at the beginning of the LS cycle.
Originality/value – This study aims to investigate whether LS can be used as a research approach by the educational research community.

Keywords Case study, Biological models, Higher order thinking skills, Lesson study as a research approach
Paper type Research paper

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 and Yoshida, 2008). Studies have shown that classrooms provide powerful, practice-based contexts in which teachers learn ways to support student learning (e.g. Opfer and Pedder, 2011). Amongst other benefits, LS has been proved to make the teachers more aware of students' thinking processes (Verhoef and Tall, 2011) and to enhance student learning (e.g. Ming Cheung and Yee Wong, 2014).

Since LS often focusses on teacher professional development, with teachers investigating their own practice, research on LS often has focussed 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 and Kihara, 2013).

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.

Funding: This work was supported by the Netherlands Organization for Scientific Research (Nederlandse Organisatie voor Wetenschappelijk Onderzoek, NWO) [grant number 023.007.065].

International Journal for Lesson & Learning Studies
Vol. 10 No. 3, 2021
pp. 286-301
Emerald Publishing Limited
2046-8253
DOI 10.1108/IJLLS-12-2020-0098

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 focusses 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 an 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 focussing on this area. LSs focus on observation of student learning may help studying students' reasoning processes. Teachers' pedagogical content knowledge 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 visualize 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 and 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 1 shows an example of a biological concept-process model

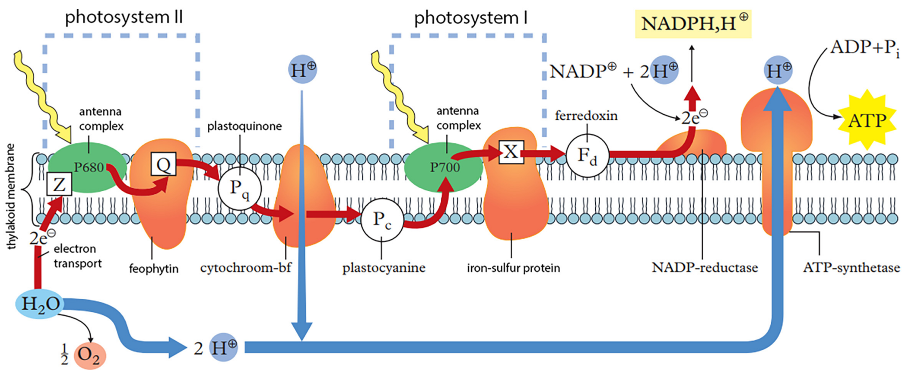


Figure 1.
A concept-process model of the light reaction of photosynthesis. Reprinted and translated with permission from Noordhoff Uitgevers, Groningen (Brouwens *et al.*, 2013)

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 not only to explain phenomena but also to formulate hypotheses or carry out thought experiments (Windschitl *et al.*, 2008).

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 1). The five aspects within this framework are as follows: *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* focusses on the extent to which the model can be compared to the original, whereas the aspect *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 and 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 focusses 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).

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 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 learnt 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:

RQ1. To what extent does the developed lesson successfully familiarize students with important levels and aspects that are associated with model-based reasoning?

RQ2. How do teachers and researchers experience using LS as a research approach?

Method

Participants

Three biology teachers (18, 30 and 9 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 an 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.

	Initial level	Level 1	Level 2	Level 3
Nature of models		(1) Model as copy (2) Model with great similarity (3) Model represents a (non-) subjective conception of the original (4) Displays a process, its components and how they are related	(1) Parts of the model are a copy (2) Model as a possible variant (3) Model as focussed representation	(1) Model as hypothetical representation
Purpose of models		(1) Model for showing the facts (2) Model for showing events	(1) Model to identify relationships	(1) Model to examine abstract/concrete ideas
Multiple models	(1) All models are the same (2) Various models of different originals (3) Only one final and correct model	(1) Different model object properties	(1) Focus on different aspects	(1) Different assumptions
Testing models	(1) No testing of models (2) Perceiving schoolbooks or their authors as authorities providing absolute truth	(1) Testing of material (2) Testing of basic requirements	(1) Comparison between original and model (2) Comparison and matching of original and model	(1) Testing hypotheses (2) Testing of hypotheses with research designs
Changing models	(1) No reason for alterations (2) Alteration of how different originals are represented	(1) Alterations to improve the model object (2) Alterations when there are errors in the model object (3) Alterations when basic requirements are not met	(1) Alterations when model does not match the original (2) Alterations due to new findings about the original	(1) Alterations due to findings from model experiments (2) Alterations when the focus of the model shifts to a different aspect of the process

Table 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)

RQ1: Pre- and post-tests

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 as follows: "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.

RQ1: 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 learnt 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 min.

RQ2: 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 min. 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 learnt from participating in an LS cycle; the extent to which they applied what they learnt to other lessons or their teaching and whether they expected to keep on using what they had learnt in the long term.

RQ2: 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-min lesson was then designed in three 2-h meetings within a time frame of two weeks. The LS team evaluated both the designed and the adapted lesson in a 1-h meeting. All meetings were audio-recorded.

Data analysis

Students' answers on the *pre- and post-tests* 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 students' 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

About 50% of the answers were coded by a second independent coder, resulting in a Cohen's kappa of 0.69 for *nature of models*, 0.87 for *purpose of models* and 0.63 for *multiple models*. Students' answers in the *audio recordings* were tagged when utterances related to aspects that were learnt 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.

Results

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 focussed on formulating key activities that let students reflect on the aspects within the framework (Table 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 the case-students, the LS team made use of the expertise of the teachers and their knowledge about the students. Since the levels in Table 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 Levels 1 and 2. Therefore, teachers were asked to define for every student whether 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 backup student was chosen and an observation scheme was formulated, in case one of the selected students would not attend the lesson. Using the observation scheme, case-students were observed by members of the LS team.

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 certain behaviour. The observer could remain seated next to this group of students and observe the backup 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 the work of every student 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 1 week later by the same teacher in a different biology class.

Lesson design – aspects of focus

The LS team decided to focus on three of the five aspects listed in Table 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 min 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 the 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*.

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 in 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 and 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).

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 focussed on the aspect *multiple models*. Students were asked to individually name differences between four models that showed the same biological process (photosynthesis) (Figure 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 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 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 3 summarizes the design process and shows the contributions of both teachers and researchers to the final lesson design.

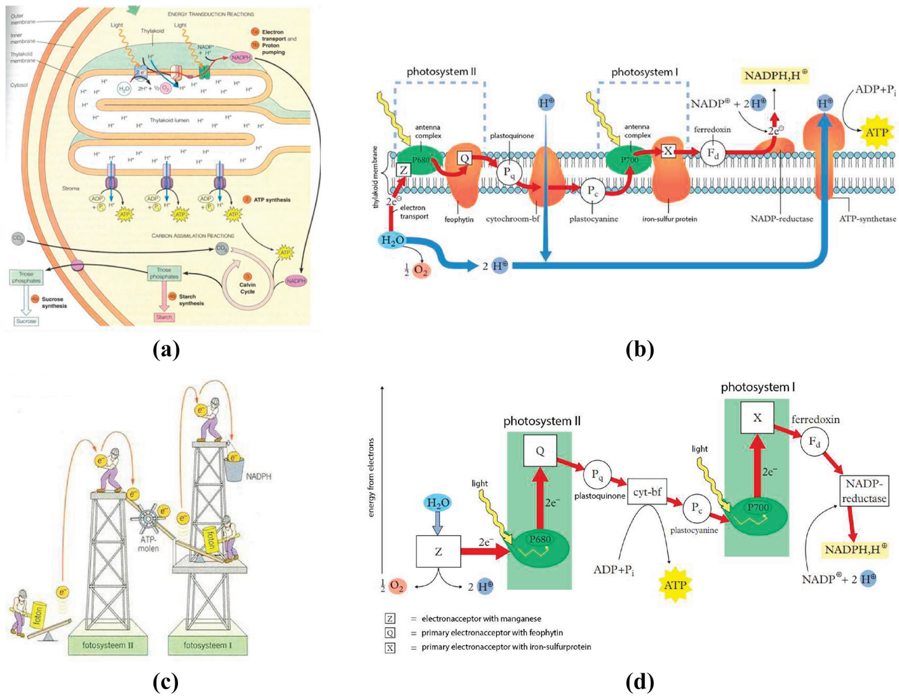
After teaching the lesson for the first time to one of the biology classes, the lesson was discussed in the LS team. Only a minor adjustment was made, the four models of photosynthesis were numbered (1–4) before teaching the adjusted lesson.

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 into 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



Note(s): All models show the same process, but have a different appearance due to differences in emphasis or choices of the creator. Figure 2a focuses on the light reaction (first part of photosynthesis), but also shows the connection to the Calvin Cycle (second part of photosynthesis). Figure 2b zooms in on a part of the thylakoid membrane with the electron transport chain, leaving out the connection to the Calvin cycle. Figure 2c focuses on the energetic state of the electrons and the role of photons in this process. Figure 2d shows resemblance with figure 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 2a and 2c) and Malmberg, 's Hertogenbosch (Figure 2b and 2d (both translated with permission from Dutch to English)) (Brouwens *et al*., 2013; Campbell & Reece, 2002)

Figure 2. The four models of the light reaction of photosynthesis that were used in the developed lesson

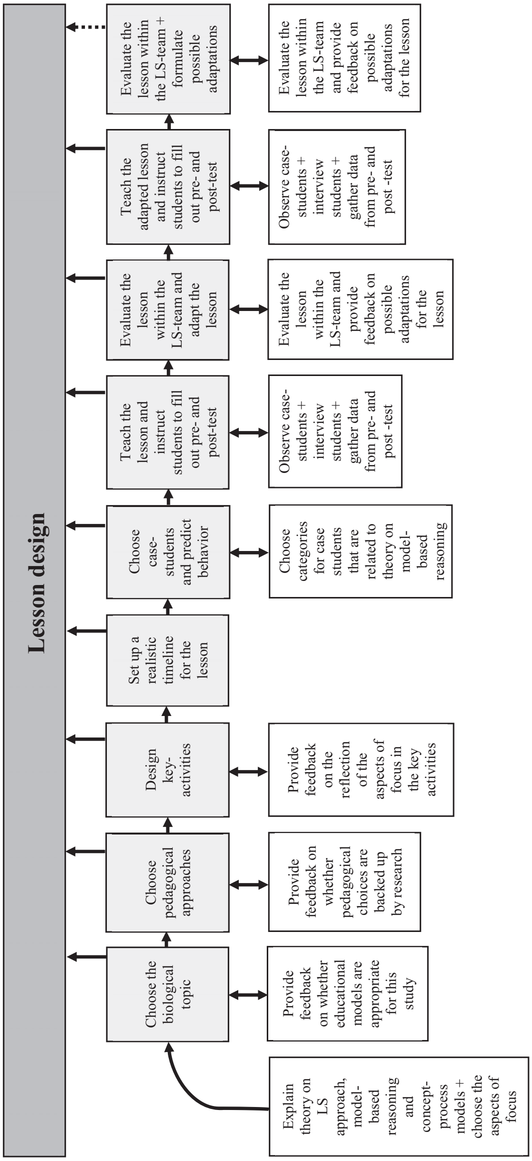


Figure 3.
Contributions of the teachers and the researchers to the lesson design. The contributions from the teachers are visualized in light grey. The contributions from the researchers are visualized in white. The arrows show interactions between the researchers, teachers and the lesson design. The 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

formulated differences related to multiple levels of reasoning within the framework (Table 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 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 aims that students were confronted with 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 2d].

L1B4: Yes, but you can see that in this model too. And in this one [pointing at the models in Figures 1a and 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 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 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 do not 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-tests, no significant differences in students' level of reasoning for each of the three aspects of focus were found. Table 2 summarizes the changes in student levels on the different aspects of model-based reasoning.

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 following quote shows how the introduction to this theory changed the teacher's view on model-based teaching, causing him or her to intend to implement this theory in his or her 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."

As 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.

	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

Note(s): 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

Table 2.
Comparison of the pre- and post-tests

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 time frame 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 would not get the point if I was not 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.

Discussion

While LS originally mainly focusses 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 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 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 the pre- and post-tests 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-tests 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-tests 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 1, the test offered the possibility to only answer on the level they preferred. Therefore, the pre- and post-tests 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 or she 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 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 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 an 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.

References

Bakker, A. (2018), *Design Research in Education: A Practical Guide for Early Career Researchers*, Routledge, Abingdon.

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

Campbell, N.A. and Reece, J.B. (2002), *Biology*, 6th ed., Pearson Education, New York, NY.

Chichibu, T. and Kihara, T. (2013), "How Japanese schools build a professional learning community by lesson study", *International Journal for Lesson and Learning Studies*, Vol. 2, pp. 12-25.

Cook, M. (2006), "Visual representations in science education: the influence of prior knowledge and cognitive load theory on instructional design principles", *Science Education*, Vol. 90, pp. 1073-1091.

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, Antwerpen.

Dudley, P., Xu, H., Vermunt, J.D. and 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*, Vol. 54, pp. 202-217.

Dudley, P. (2015), *Lesson Study, Professional Learning for Our Time*, Routledge, Abingdon.

Fernandez, C. and Yoshida, M. (2008), *Lesson Study: A Japanese Approach to Improving Mathematics Teaching and Learning*, Lawrence Erlbaum Associates, Mahwah, New Jersey.

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*, Vol. 3, p. 16.

Grosslight, L., Unger, C., Jay, E. and 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*, Vol. 28, pp. 799-822.

Grünkorn, J., Upmeier zu Belzen, A. and Krüger, D. (2014), "Assessing students' understandings of biological models and their use in science to evaluate a theoretical framework", *International Journal of Science Education*, Vol. 36, pp. 1651-1684.

Harrison, A.G. and Treagust, D.F. (2000), "A typology of school science models", *International Journal of Science Education*, Vol. 22, pp. 1011-1026.

Jansen, S., Knippels, M.C.P.J. and van Joolingen, W.R. (2019), "Assessing students' understanding of models of biological processes: a revised framework", *International Journal of Science Education*, Vol. 41, pp. 981-994.

Ming Cheung, W. and Yee Wong, W. (2014), "Does lesson study work?", *International Journal for Lesson and Learning Studies*, Vol. 3, pp. 137-149.

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*, Vol. 37, pp. 353-369.

Opfer, V.D. and Pedder, D. (2011), "Conceptualizing teacher professional learning", *Review of Educational Research*, Vol. 81, pp. 376-407.

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*, Vol. 68, pp. 289-303.

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 questions", *Science*, Vol. 323, pp. 122-124.

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*, Vol. 8, pp. 89-94.

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

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)*, Ankara, Turkey, 10-15 July 2011.

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*, Vol. 81, pp. 61-73.

Windschitl, M., Thompson, J. and Braaten, M. (2008), "Beyond the scientific method: model-based inquiry as a new paradigm of preference for school science investigations", *Science Education*, Vol. 92, pp. 941-967.

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*, Vol. 100, 101534.

Supplementary material

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

(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

- The existence of both of these models is important
- One of the models is better/more useful than the other
- 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 focussing on the binding of the anticodon on tRNA to the codon on mRNA, and one model focussing 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

Corresponding author
Wouter R. van Joolingen can be contacted at: w.r.vanjoolingen@uu.nl