On the influence of knowledge about the ideal-typical modelling processes on individuals' modelling routes

Sobre a influência do conhecimento acerca dos processos ideais-típicos de modelação nas rotas de modelação dos indivíduos

Lisa Schneider
University of Kaiserslautern
Germany
lisa.schneider@mathematik.uni-kl.de

Rita Borromeo Ferri
University of Kassel
Germany
borromeo@mathematik.uni-kassel.de

Stefan Ruzika
University of Kaiserslautern
Germany
ruzika@mathematik.uni-kl.de

Abstract. Working on mathematical modelling tasks is challenging for students. Several studies have shown that knowledge of mathematical modelling on a meta-level has a positive effect on the modelling process. Nevertheless, students mostly do not knowingly and consciously use solution strategies when working on modelling tasks. Within the framework of our study, we investigated whether and to what extent knowledge about ideal-typical modelling processes has an effect on the structure of the solution processes of individuals. Individuals acquired this knowledge in our study in the form of an instruction that includes information about the modelling process, e.g., the modelling cycle and a solution plan. In this article, the structure of individual modelling routes of students who have received an instruction about modelling processes are compared with those students without such an instruction. The data in the study was collected, presented, and analysed using the Modelling-Activity-Interaction-Tool (MAI-Tool), which is also presented here. The MAI-Tool is a newly developed instrument based on quantitative methods to capture and analyse structures and patterns of modelling processes in more detail than with previously known methods.
Keywords: individual modelling routes; structure of modelling processes; knowledge about modelling processes; MAI-Tool.

Quadrante: Revista de Investigação em Educação Matemática
https://doi.org/10.48489/quadrante.23719
Received: February 2021 / Accepted: November 2021

Resumo. Trabalhar em tarefas de modelação matemática é um desafio para os estudantes. Vários estudos demonstraram que o conhecimento sobre a modelação matemática, a um meta-nível, tem um efeito positivo sobre o processo de modelação. No entanto, os estudantes não utilizam estratégias de resolução, intencional e conscientemente, ao trabalharem em tarefas de modelação. No âmbito do nosso estudo, pretende-se saber se, e em que medida, o conhecimento sobre os processos ideais-típicos de modelação tem um efeito sobre a estrutura dos processos de resolução dos indivíduos. Os indivíduos adquiriram esse conhecimento, durante o nosso estudo, no contexto de um ensino que incluiu informação sobre o processo de modelação, tal como, por exemplo, o ciclo de modelação e um plano de resolução. Neste artigo, a estrutura das rotas de modelação individuais dos estudantes que receberam instrução sobre os processos de modelação é comparada com a dos estudantes que não receberam tal instrução. Os dados do estudo foram recolhidos, apresentados e analisados, utilizando a Ferramenta de Modelação-Atividade-Interação (MAI-Tool), que também é aqui apresentada. O MAI-Tool é uma ferramenta recentemente desenvolvida com base em métodos quantitativos para captar e analisar estruturas e padrões de processos de modelação com mais detalhe do que com métodos previamente conhecidos.
Palavras-chave: rotas de modelação individuais; estrutura dos processos de modelação; conhecimentos sobre processos de modelação; MAI-Tool.

Introduction and research question

The term mathematical modelling is understood as the process of solving real-world problems with the help of mathematical methods (Greefrath et al., 2013). The mathematical modelling process thus presupposes an intensive processing and solving of the problem. Problems from reality (the extra-mathematical world) are transferred to mathematics in a simplified form and, after being solved in mathematics, are returned to reality (Niss et al. 2007). A modelling process usually consists of several turns of different steps, which can be idealised in a modelling cycle. The steps in the modelling process, also called phases, follow a fixed order. The transitions between the phases are described with the help of activities. Modelling cycles are used to illustrate the concept of mathematical modelling, to support learners in working on modelling tasks, especially in the area of metacognition, and to form the basis for empirical studies (Blum, 1996).

In order to investigate modelling processes, the modelling cycle is often used as an analysis tool. With the help of a modelling cycle, individual modelling processes can be reconstructed and, on the basis of this, the process can be represented and analysed (Borromeo Ferri, 2010).

The Modelling-Activity-Interaction-Tool (MAI-Tool) was developed to collect, represent, and finally evaluate modelling processes (Ruzika & Schneider, 2020). The modelling process is digitally recorded according to the phases of the modelling cycle. The evaluation is therefore based on the structure of the modelling cycle, e.g., phases and their transitions as well as patterns. The analysis can thus not only be carried out qualitatively – as before – but also quantitatively since modelling processes are described with numerical data (see Section The Modelling-Activity-Interaction-Tool (MAI-Tool)). The quantitative evaluation gives us a more detailed insight into modelling processes, as it complements the qualitative evaluation.

Several empirical studies have already shown that knowledge about the modelling process has a positive influence on the modelling process (Schukajlow et al., 2015; Stillman & Galbraith, 1998). This knowledge about mathematical modelling is usually communicated to students as a solution plan (Beckschulte, 2019; Schukajlow et al., 2015; Zöttl et al., 2011). In previous studies it was investigated whether students with a solution plan possessed more modelling skills than those without. With regard to the effectiveness of solution plans, empirical research has not yet been conducted to determine "whether the solution plan plays a role in the solution process or only influences the outcome" (Greefrath, 2018, p. 46).

In an empirical study, we investigated individuals' solution processes when working on a modelling task by comparing the individual modelling routes of students who were instructed about modelling processes with those who were not. The knowledge about modelling processes imparted to the individuals is given to them as an instruction. Thus, we neither analyse the influence on the result nor the differences in the modelling competencies, but we focus on the structure of the solution processes. In this article, we examine the individual modelling routes for differences in the number and duration of the phases that occur over total time during the modelling process. To make the values comparable, we choose the relative frequency in each case, i.e., we compare the relative number as well as the relative duration of each phase in the modelling process from individuals with and without instruction.

Students aged 15-16 (10th grade of German High School – *Gymnasium*) participated in the study. The study was conducted as part of the mathematics lessons at school. In a group size of five, students worked on the "Filling-up" task (Blum & Leiss, 2006), for which they had about 30 minutes to solve. A total of 40 students were selected for the study. Before the students worked on the task, they were divided into two cohorts: one half received an instruction about ideal-typical modelling processes, the other did not.

The instruction was based on the meta-level according to the guidelines of Borromeo Ferri (2018) and Vorhölter and Kaiser (2016). The instruction includes knowledge about the course of the ideal-typical modelling process in the form of a modelling cycle and the related solution plan. For this study, we have chosen the five-step modelling cycle of Kaiser and Stender (2013) (Figure 1), which, in addition to the instruction, also forms the basis for the evaluation of the individual modelling routes. The respective phases of the modelling cycle can be found in the five-step solution plan by Beckschulte (2019). Linking this modelling cycle and this solution plan is suitable for an instruction because the modelling process can be represented graphically and also expressed verbally. In addition, a prototypical modelling problem was explained using the modelling cycle as an example: when colouring the map of Germany, it is relevant for the formation of the real model which federal states have a common border (Leuders, 2007). This is modelled in a graph where the nodes represent the federal states and a common border is represented by a common edge. By applying a short algorithm, the nodes are coloured so that the colour of the node corresponds to the colour of the federal state. This real solution is validated: a decision is made whether the real model can be improved. This modelling problem was chosen because it can be solved without mathematical formulae, so that the focus is on the steps of the modelling cycle. The instruction was always given by one of the researchers. The students were involved in the task by means of clarifying questions. In total, the instruction lasted about 20 minutes. While working on the modelling task, the modelling cycle was visible for the students.

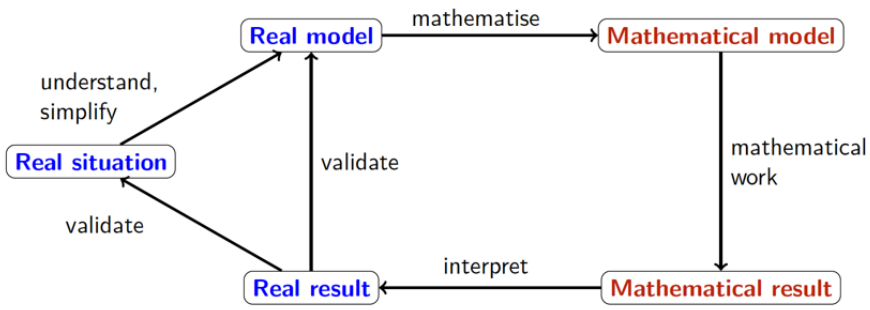


Figure 1. The modelling cycle (Kaiser & Stender, 2013)

In the following, we will call the students who received such an instruction the *instructed students* and the students who did not receive an instruction the *non-instructed students*. Furthermore, we call the knowledge about modelling processes imparted to the individuals in the instruction *meta-knowledge*. However, we do not check to what extent they actually acquired this.

We investigate whether instructed students proceeded differently when working on a modelling task than non-instructed students. Thus, the central research question in this article is:

> Does the knowledge about ideal-typical modelling processes obtained from an instruction influence the solution process structure of individuals when working on a modelling task, in particular, to what extent do the relative number and relative duration of the phases of instructed students differ from the non-instructed students?

After presenting the theoretical background in the following section, the methodology as well as the study design are described. Then, the focus is on the results of the study and their interpretation. Finally, the last section discusses further research.

Theoretical framework

In this section, central elements of the study's theoretical framework are presented. The first subsection focuses on existing methods for collecting and analysing modelling processes. A new approach presented in the second subsection is the MAI-Tool. The third subsection treats the theory of meta-knowledge about mathematical modelling. Hypotheses are formulated in the fourth subsection, whose verification and interpretation are explained in section Results of the study.

Analysis of individual modelling routes

Methods and concepts that are suitable for the representation and analysis of modelling processes already exist: the solution processes when working on a modelling task can be analysed at the level of the individuals. These processes can be described and characterised with the help of ideal-typical modelling cycles.

Empirical studies have already examined modelling processes when solving a modelling task (Borromeo Ferri, 2011; Matos & Carreira, 1997). It was found out that the modelling process – based on the modelling cycle – is not linear and switches back and forth between the phases. Borromeo Ferri (2007, p. 265) calls this an individual modelling route: "The individual starts this process in a certain phase . . . and then goes through different phases several times or only once". It can also happen that some phases are omitted, while others are more pronounced. The individual modelling routes are "visible modelling routes" (Borromeo Ferri, 2007, p. 265), i.e., they are based on verbal statements and visible actions by the individuals.

Individual routes are visualised on the modelling cycle by connecting successive phases with arrows (Figure 2).

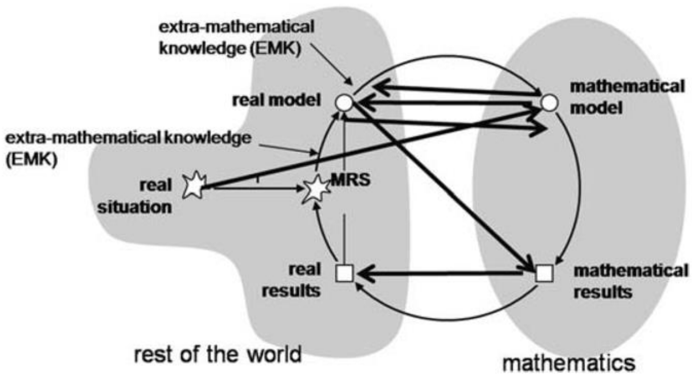


Figure 2. An individual modelling route by Borromeo Ferri (2010, p. 113)

Thus, the modelling cycle is not only used for illustration, but also as an analysis tool. This representation can be used to reconstruct the structure in individual routes. With the help of these patterns, the processes can be categorised and by means of this representation, a structural comparison of individual modelling processes is possible.

Another concept for representing and analysing modelling processes is the Modelling-Activity-Diagram (MAD) by Ärlebäck and Albarracin (2019) that captures activities instead of modelling phases. The activities describe what is to be done in the phases of a modelling cycle, e.g., the activity calculating refers to the *mathematical model*. The MAD gives a linear representation of the activities during modelling over time (Figure 3) and can be applied to a group but also to a single individual.

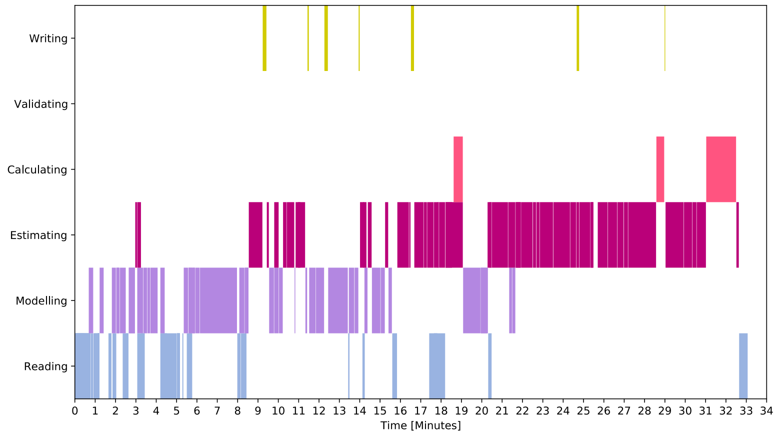


Figure 3. The Modelling-Activity-Diagram (MAD) (created with the MAI-Tool)

In both approaches, the data is interpreted using qualitative methods. Thus, results of a study are described using excerpts as examples. However, in order to investigate modelling processes in a more precise and a generalised way, the structure of the modelling processes should also be described with numerical data using a quantitative investigation instrument, so that an evaluation with statistical tests is also possible. A digital recording of the modelling processes facilitates its evaluation with numerical data.

The Modelling-Activity-Interaction-Tool (MAI-Tool)

The concepts of Borromeo Ferri (2011) and Ärlebäck and Albarracin (2019) for the analysis of modelling processes are incorporated and extended in a new research instrument: in the MAI-Tool, modelling processes are digitally recorded and evaluated using implemented algorithms (Ruzika & Schneider, 2020). Numerical data are determined from the qualitative data collection and allow a quantitative evaluation with statistical tests.

The concept behind the MAI-Tool is based on interactions within the group while they work on the modelling task. An interaction is understood as the interdependence of the behaviour of two or more individuals, i.e., individuals refer to verbal or non-verbal acts of other group members (Kolbe & Boos, 2018). Interaction units are entered into the MAI-Tool and contain the following information: interacting person(s), duration as well as content of the interaction. A timestamp is added automatically. Interaction units can consist of several acts of group members. This is the case when individuals react to each other through verbal or non-verbal communication, i.e., the acts are interdependent in terms of content. For example, the following acts are combined into one interaction unit. The phase *real model* is assigned to this interaction unit.

Student 1: If Mrs Stein refuels in Luxembourg, then we have to take into account the petrol consumed for the trip.
Student 2: Does anyone know what the fuel consumption of a car is?
Student 3: Let's assume that the car consumes seven litres per 100 km.

Furthermore, we consider an act or acts of a single individual as an interaction, even if they do not depend on any act of another group member.

Student 1: *Underlines text passages on the task sheet.*
Student 2: The variable x is the amount of petrol in litres and $T(x)$ is then the cost of refuelling in Trier.
Student 3: Then let us now formulate the equation.

The act of student 1 is in an interaction unit, while the acts of student 2 and student 3 are combined into one interaction unit, because student 1 is in the real model, but student 2 and student 3 are in the mathematical model. An interaction unit consists only of the acts that are interdependent, i.e., that follow each other in time and are assigned to the same phase.

Interaction units are entered into the MAI-Tool; improvements to the entries are possible ad-hoc as well as retrospectively (Figure 4).

Figure 4. Input and improvement of the data in the MAI-Tool

To finally present and evaluate the data, a mathematical model is used, on the basis of which the modelling process of the group over time is structurally understood as a dynamic social network. This model contains the information about who interacts with whom in which phase of the modelling cycle and for how long. Here, the nodes of the network represent the individuals and the edges, which appear and disappear over time, represent the interaction. Each interaction is assigned to a phase of the modelling process: each edge is coloured with the corresponding colour of the phase (Figure 5).

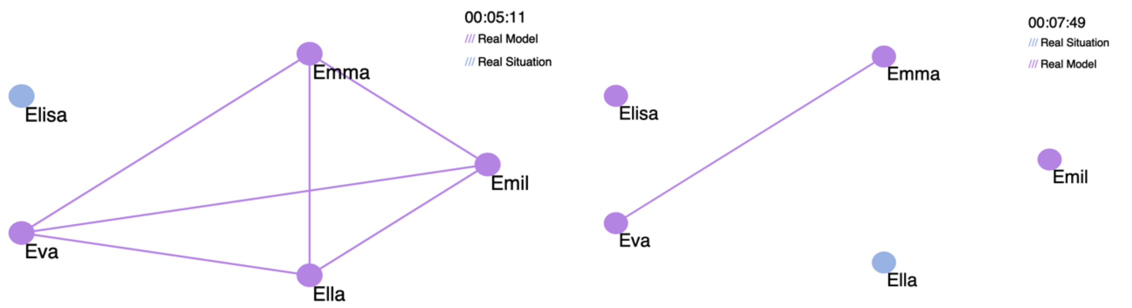


Figure 5. Excerpts from the Modelling-Activity-Interaction-Network (MAIN)

The modelling process can be represented at any point in time. Ruzika & Schneider (2020) call the network Modelling-Activity-Interaction-Network (MAIN).

The evaluation of the individual modelling routes is based on the MAIN, i.e., data are extracted from the MAIN and automatically prepared for each individual. In addition, it is possible to select which properties of the modelling process are to be evaluated, e.g., phase transitions, phases, blocks, and patterns (Figure 6). Section *Data collection with the MAI-Tool* presents and explains the evaluation features for this study.

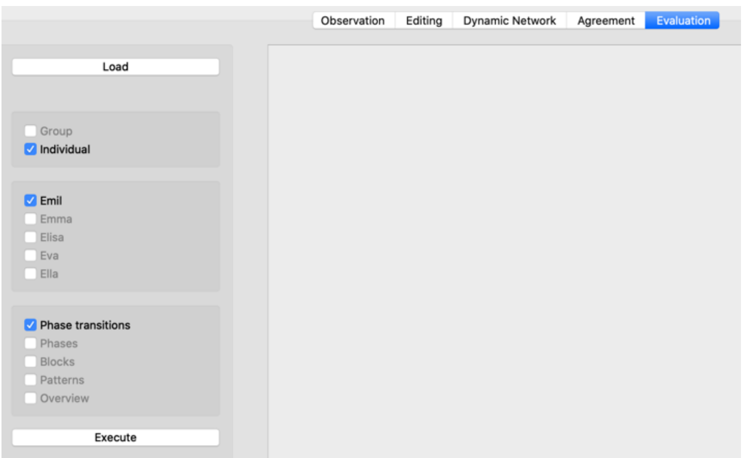


Figure 6. Evaluation mask in the MAI-Tool

Meta-knowledge about the mathematical modelling process

Dealing with modelling problems in the form of mathematical modelling tasks is complex and therefore challenging for students (Blum, 2015). As the solution process of such a task consists of several steps, each sub-step of the modelling process can represent a cognitive hurdle (Blum, 2015; Galbraith & Stillman, 2006): students are challenged to develop their own solution approaches that include mathematical procedures in addition to deciding which assumptions are relevant for developing a solution.

Several studies have shown that students generally do not follow a conscious strategy when solving modelling tasks and they often do not know how to proceed when difficulties arise (Blum, 2015; Kaiser et al., 2015). In particular, a lack of metacognitive skills can cause problems when working on a modelling task. Metacognitive modelling competencies can be divided – based on the theories of metacognition – into declarative meta-knowledge and procedural metacognitive strategies (Flavell, 1979; Vorhölter et al., 2019). The first group includes knowledge about features of modelling tasks, solution strategies, and knowledge about one's own abilities and those of group members. Students are taught to choose their own solution approach by making simplified assumptions to develop the corresponding model. Different approaches are possible. The contextual knowledge presented in the task is not sufficient to solve the task, so assumptions have to be made and extra-mathematical knowledge has to be included. The modelling cycle, for example, is suitable for teaching strategies for solving modelling tasks. The individual steps of the modelling process can be illustrated and explained using this cycle. The use of metacognitive knowledge is judged to be increasingly relevant for processing a modelling task, as several studies point out (Stillman, 2011). Therefore, there is a demand to apply strategies to overcome cognitive hurdles in the processing of modelling tasks (Stillman, 2004). But which measures can be taken to support students in solving a modelling task? When using the modelling cycle as a solution plan, the focus is on the sub-processes, i.e., the phases. A solution plan usually consists of the presentation of a modelling cycle, where each phase or each transition between the phases is described in more detail by means of activities. These are supported with questions or prompts (Brand & Vorhölter, 2018). Such a solution plan belongs to the indirect general strategic aids, as there are no concrete aids related to the problem of the task (Borromeo Ferri, 2006). Greefrath (2014) has presented several variants of solution plans, which differ in the number as well as the type of steps, among other things. The four-step solution plan, which can be found in Blum (2007), comprises the sub-steps *understanding the task*, *creating the model*, *using the mathematics, and explaining the result*. Zöttl et al. (2011) limit their solution plan to three steps and neglect the step *using the mathematics* in comparison to the four-step solution plan. Beckschulte (2019) developed a five-step solution plan in which the step of *validating the result and the solution* is emphasised.

By presenting the entire solution process in the form of the solution plan, it is possible to acquire meta-knowledge about the processing of modelling tasks: "[The] solution plan is not meant as a schema that has to be used by the students' but as an aid for difficulties that may occur in the course of the solution process" (Blum & Borromeo Ferri, 2009, p. 55). This information should be known to the students before they start working on the task, as in this way possible problems can be anticipated and reduced. Results of empirical studies have also already shown that the use of solution plans serves as an orientation aid for the students (Maaß, 2004): hurdles in the respective phases of the modelling process are reduced.

Hypotheses of the study

The aim of the empirical study is to find out whether and to what extent the solution processes of the instructed students differ from those of the non-instructed students when working on a modelling task and how this is noticeable in the structure of the modelling processes. We want to examine this by testing hypotheses. The hypotheses are based on the modelling cycle of Kaiser and Stender (Figure 1).

The hypotheses test the extent to which the relative number and relative duration of a phase aggregated over the total time of the modelling process differs in the two cohorts. By the relative number of a phase we mean the relative frequency of the occurrence of that phase in the modelling process. Similarly, the relative duration indicates the relative frequency in relation to the duration of a phase in the entire modelling process. There are expected differences in the occurrence and duration of certain phases in the modelling processes, which we formulate below.

Non-instructed students do not take much time to analyse the initial problem in contrast to the instructed students. The latter take more time to understand the modelling task and accordingly, they also switch more often to the phase *real situation*.

- (H1) The relative number of times that individuals switch to the phase *real situation* as well as its relative duration are significantly higher for the instructed than for the non-instructed students.

Without any knowledge of the ideal-typical modelling process, individuals have difficulties developing a mathematical model and consequently difficulties solving it. Instructed students should therefore spend more time and switch more often to the phases *mathematical model* and *mathematical solution* during the modelling process. This means for the relative number and relative duration of the phases:

- (H2) The relative number of times that individuals switch to the phase *mathematical model* as well as its relative duration are significantly higher for the instructed than for the non-instructed students.
- (H3) The relative number of times that individuals switch to the phase *mathematical solution* as well as its relative duration are significantly higher for the instructed than for the non-instructed students.

The modelling cycle can be divided into the *real world* and the *mathematical world* (Pollak, 1979). The *real world* contains phases that can be assigned to reality: *real situation, real model, real solution*. Thus, the *mathematical world* includes the *mathematical model* and the *mathematical result*. Derived from the hypotheses (H2) and (H3), it can be assumed that instructed students focus more on the *mathematical world* as they spend more time and longer in the phases *mathematical model* and the *mathematical result*:

- (H4) The relative number and the relative duration of the phases taking place in the *mathematical world* are significantly higher for the instructed than for the non-instructed students.

Methodology and design of the study

The following section presents the methodology and the next the study design. This study follows a mixed-methods approach. The data collection and documentation are qualitative, while the data is analysed quantitatively with a statistical test.

Methodology

The investigation of structures in individual modelling processes requires a qualitative approach. Since the influence of an instruction about the modelling process is also being investigated, i.e., the extent to which solution processes of instructed students differ from the non-instructed students, a quantitative approach is also necessary. Therefore, a mixed-methods approach is pursued in this study, in which qualitative and quantitative procedures are combined in order to achieve the highest possible gain in knowledge (Döring & Bortz, 2016).

The combination of both research approaches is necessary for this study. The students' individual modelling routes are reconstructed and documented on the basis of observation. This is only possible if the modelling processes are recorded and subsequently collected in detail. Consequently, the collection and documentation are qualitative. Each individual modelling process is documented with regard to frequencies concerning certain characteristics to be investigated, in order to then evaluate it statistically on the basis of these facts. Therefore, the evaluation is assigned to the quantitative approach.

The data documentation is done qualitatively according to the Grounded Theory (Strauß & Corbin, 1996). The central process is coding. In the construction of the coding scheme, the underlying ideal-typical modelling cycle by Kaiser and Stender (Figure 1), including the phase descriptions, was used as a framework. This knowledge and the individual terms of the phases are used to develop theoretical codes. These are complemented by in-vivo codes added to the coding scheme during the coding process. Each unit of interaction is assigned a theoretical code as well as the names of the people involved.

The inputs are automatically processed and evaluated in the MAI-Tool for each individual. Then the data is statistically analysed. Since a small sample size ($N=40$) is examined in the study, a non-parametric procedure is needed, which, in contrast to t-tests, requires fewer prerequisites. Only the prerequisites that the samples are independent and the dependent variable (the respective characteristic to be examined) is ordinally scaled must be fulfilled, which is, in fact, the case. Since the two samples are independent, significant differences between the instructed and non-instructed students can be detected with the Mann-Whitney-U-test. A prerequisite for a significant result is that the *p*-value for the tested hypothesis is smaller than *0.05*. The two samples are compared in terms of their mean values (of the dependent variables). Since we test directed hypotheses, i.e., whether one cohort has a higher or lower mean value than the other, one-sided testing is carried out.

Design of the study

The sample of the study comprises 40 students aged 15-16 from four 10th grade classes from different High Schools. The study was conducted at school during mathematics lessons. Each class was divided into a cohort that received an instruction about modelling processes and a cohort that did not receive such an instruction. For this purpose, the cohorts were placed in different rooms.

Before groups were formed in each cohort to work on the modelling task, each student was given a questionnaire to test the knowledge about mathematical modelling. This questionnaire includes questions about whether the students have worked on a modelling task before, e.g., in regular classes, in school projects or in extracurricular projects. In addition, it was asked whether the students know how the process of solving a modelling task works, i.e., whether they have already gained knowledge about the procedure of a modelling process at the meta-level. Although modelling is required as a competence in the curriculum, it is treated differently by each teacher in the mathematics lessons. Within each cohort, groups of five students were formed based on the results of the questionnaire and one group participated in the study. These students indicated in the questionnaire that they had not previously worked on a modelling task or received knowledge about the modelling process. By selecting only those students without modelling experience for the study, the influence of previous experience was to be excluded. In total, four groups with instruction and four groups without instruction participated, i.e., $n_1=20$ instructed students and $n_2=20$ non-instructed students.

Each group within the cohorts completed the “Filling-up” task (Blum & Leiß, 2006), regardless of whether they participated in the study or not. The task is about the question whether it is profitable for a certain Mrs. Stein to drive from her home town, Trier, across the border to Luxembourg in order to fill up her tank there because the petrol is cheaper. The groups had about 30 minutes to work on the task, although they were given a few extra minutes if they were close to finding a solution. The groups that took part in the study were videorecorded while working on the problem. The groups – regardless of cohort and participation in the study – did not receive any additional assistance during the processing. Thus, the solution processes are not influenced by help from the teachers.

In the study, all groups worked on the same task, so that a possible different solution process is not due to the respective modelling task. Thus, the same framework conditions were created for all groups, so that there were no differences within the instructed and non-instructed groups. Figure 7 presents an overview of the study design.

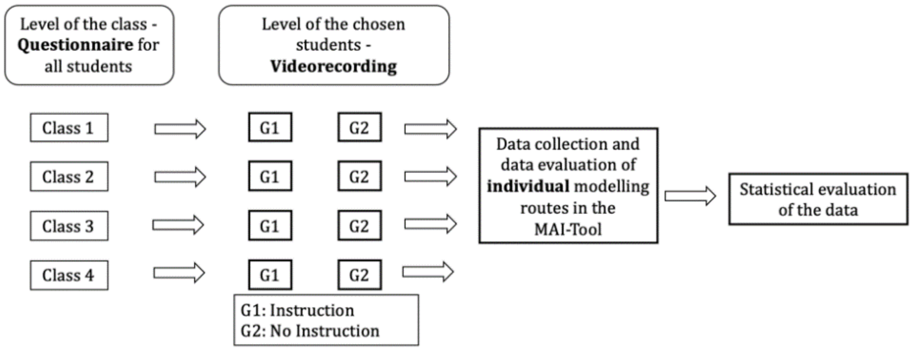


Figure 7. Design of the study

Data collection with the MAI-Tool

Since the focus in the study is on the structure, we concentrate on verbal and non-verbal communication in the form of interactions that can be observed when documenting the modelling processes. Cognitive processes are not recorded, as they are not relevant for the structure.

The coding scheme was created on the basis of the videorecorded groups. As already mentioned, the coding procedure is applied in the sense of Grounded Theory according to Strauß and Corbin (1996). The modelling process is divided into coding units, where the following information is documented: who is interacting (with whom)? In which phase of the modelling process are the individuals/the individual? When does the coding unit start and end? A coding unit can consist of one or more individuals. Each new interaction unit that occurs is documented, i.e., it is event-driven coding and not at fixed points in time.

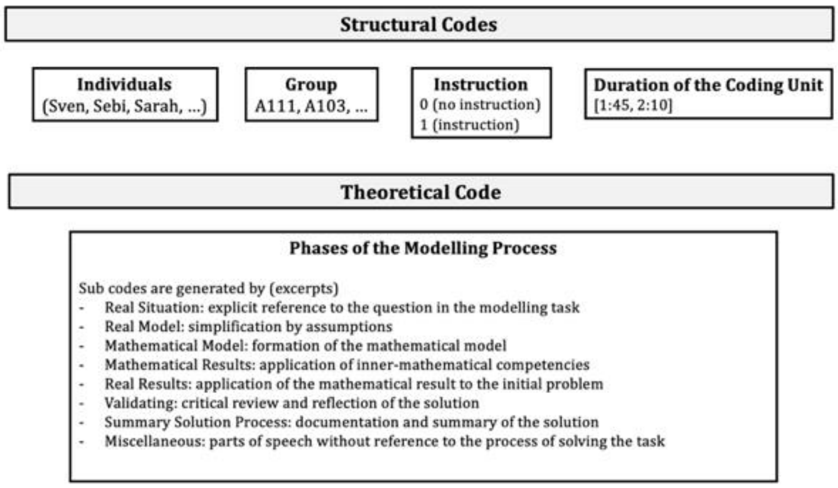


Figure 8. The coding scheme

In the coding guide, it is excluded that an individual is involved in two simultaneous coding units, as this facilitates the evaluation. A new coding unit is chosen if there is a change of speaker that does not relate (in terms of content) to the current coding unit or if an individual move to another phase. A coding unit consists of structural codes and a theoretical code (Figure 8):

- structural codes: individuals, group, instruction (yes/no), duration of the coding unit
- theoretical code: phases of the modelling process.

In addition to the phases of the modelling cycle, in-vivo codes have been added to the coding scheme: checking the plausibility of the solution cannot be clearly assigned to a phase, so *validating* is added as a sub code. If the group recapitulates, documents or summarises the solution process, this is coded with the sub code *summary solution process*. Interactions that are not thematically concerned with the processing of the modelling task are coded with *miscellaneous*. *Miscellaneous* includes off-task speech as well as off-task work and is coded exactly when the interaction is clearly not related to the task or when no other sub-activity (e.g., entering into the calculator) is observed. In the further course of the article, the term phase includes in addition to the phases of the modelling cycle the sub codes *validating*, *summary solution process* and *miscellaneous*. The coding guide defines when an interaction is assigned to which phase. Figure 8 briefly describes each phase. After the modelling process has been coded and entered into the MAI-Tool in a first run, the inputs can be checked and improved in the “edit mode” (Figure 4).

In order to ensure reliability for the recording of the modelling processes as well as for the coding, the data were coded by two persons. The second coder received training so that all observations were coded under comparable conditions. Cohen’s Kappa (Cohen, 1960) was used to check the interrater reliability for each individual, which ranged from κ= [.72, .90]. Since the time component also plays a role in checking the interrater reliability, the agreement of the theoretical code was checked for each second.

Before the modelling processes are statistically evaluated, the data are automatically output in the MAI-Tool for each individual. The phases are evaluated according to their number of occurrences and their duration. Table 1 shows a concrete example: on the one hand, the absolute and relative frequency of how often an individual was in a given phase is calculated. The temporal duration, summed up over the total time of the modelling process, is also calculated for each phase. Here, too, the absolute and relative frequencies are calculated. Looking at the phases in terms of occurrence as well as duration provides a complete view of the phases: even if the number of occurrences of a phase is high, this does not necessarily mean that the temporal duration is also high.

Table 1. Excerpts from the evaluation in the MAI-Tool: the occurrence (left) and duration (right) of the phases for an individual in the modelling

Phases	Absolute frequency (occurrence)	Relative frequency (occurrence)	Total duration	Relative duration
real situation	6	6.7%	01:09	3.7%
real model	25	28.1%	06:47	21.7%
mathematical model	10	11.2%	02:35	8.3%
mathematical solution	4	4.5%	00:30	1.6%
real result	6	6.7%	01:42	5.4%
validating	2	2.2%	00:30	1.6%
summary solution process	18	20.2%	07:53	25.2%
miscellaneous	18	20.2%	05:59	19.1%

Statistical evaluation of the data

As explained before, the Mann-Whitney-U-test is applied to detect significant differences between the instructed and non-instructed students. The significance level is set at $\alpha=0.05$. The aim of the statistical test is that the null hypothesis is rejected, because then the test gives a significant result. Therefore, the hypotheses formulated are used as alternative hypotheses.

Afterwards, the effect size of the test is computed, which enables an exact specification of the alternative hypothesis and is calculated depending on the statistical test. The effect size δ is calculated according to Fritz et al. (2012) and interpreted according to Cohen (1988). The Mann-Whitney-U-test and the calculation of the effect size for this study are carried out with the use of the software R.

Since we are testing multiple hypotheses from one data set, these analyses are corrected using the Benjamini-Hochberg method to control for multiple comparisons and false discovery rate (FDR) (Benjamini & Hochberg, 1995).

Results of the study

In this section we present the results of the study. The first subsection compares the instructed and the non-instructed students. The results are then interpreted in the second subsection.

Comparison of individual modelling routes with and without instruction about ideal-typical modelling processes

The hypotheses (H1)-(H4) are tested using the Mann-Whitney-U-test, the results of which are shown in Table 2. The p-values for the respective variables are always below the significance level of 5% ($\alpha=0.05$). From this we conclude that all hypotheses that correspond to the alternative hypothesis in the Mann-Whitney-U-test are accepted. For each hypothesis, the effect size is at least weak. The values of the effect sizes for each hypothesis underline that an effect is measured.

Table 2. Results of the Mann-Whitney-U-test for group 0 (non-instructed students) and group 1 (instructed students) with p-values and effect size δ (rounded to 3 decimal positions)

Variable (corresponding hypothesis)	Mean value Group 0	Standard deviation Group 0	Mean value Group 1	Standard deviation Group 1	p-value	Effect size δ
relative number *real situation* (H1)	16.3%	0.077	22.1%	0.118	0.042	0.242
relative duration *real situation* (H1)	9.4%	0.042	13.4%	0.057	0.01	0.328
relative number *mathematical model* (H2)	16.8%	0.142	22.5%	0.110	0.019	0.289
relative duration *mathematical model* (H2)	12.6%	0.133	17.8%	0.109	0.032	0.259
relative number *mathematical result* (H3)	1.2%	0.016	3.8%	0.024	< 0.001	0.459
relative duration *mathematical result* (H3)	0.6%	0.008	3.6%	0.025	<0.001	0.51
relative number *mathematical world* (H4)	18%	0.145	26.2%	0.112	0.019	0.292
relative duration *mathematical world* (H4)	13.1%	0.138	21.4%	0.127	0.011	0.322

During the evaluation, a new, previously unexplored phenomenon occurred, which we describe in more detail. Parts of speech that did not refer to the processing or the solution process of the modelling task were coded with *miscellaneous*. This was added to the coding scheme during the coding process because a huge number of such parts of speech were observed. The mean values for relative number and relative duration of *miscellaneous* for the non-instructed students are higher than those of the instructed students. The application of the Mann-Whitney-U-test as well as the calculation of the effect size show that instructed and non-instructed students differ significantly. More precisely, the relative number as well as the relative duration of *miscellaneous* are significantly higher for the non-instructed students than for the instructed students (Table 3).

The mean values of both cohorts also differ in the relative number and the relative duration of the phase *summary solution process*: The mean values of the non-instructed students are higher than those of the instructed students. Here, too, the application of the Mann-Whitney-U-test showed this difference to be significant.

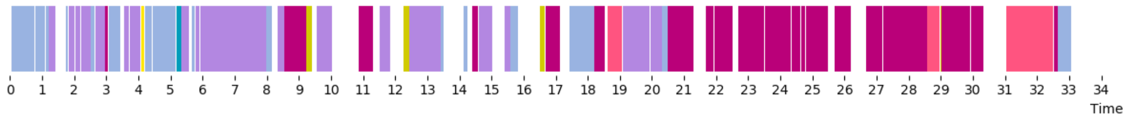
Table 3. Results of the Mann-Whitney-U-test for group 0 (non-instructed students) and group 1 (instructed students) with p-values and effect size δ (rounded to 3 decimal positions) for *summary solution process* and *miscellaneous*

Variable	Mean value Group 0	Standard deviation Group 0	Mean value Group 1	Standard deviation Group 1	p-value	Effect size δ
relative number *summary solution process*	7.8%	0.080	1.2%	0.013	0.029	0.267
relative duration *summary solution process*	7%	0.078	0.08%	0.01	0.022	0.28
relative number *miscellaneous*	16.7%	0.134	7.3%	0.083	0.004	0.36
relative duration *miscellaneous*	14.8%	0.113	4.1%	0.053	0.001	0.434

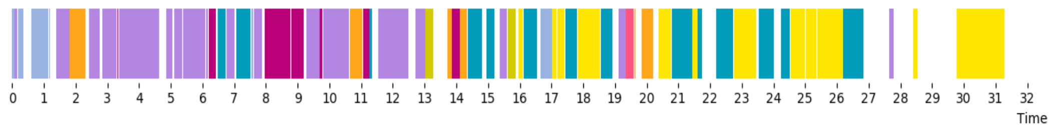
Interpretation of the results

The results of this study show that a deeper analysis becomes possible with the MAI-Tool. After a qualitative input of the data, algorithms reproduce it quantitatively in numerical data. The evaluation is ad-hoc, as the data is available digitally.

Testing the hypotheses with the Mann-Whitney-U-test proved the validity of the statements. The results are illustrated by two typical cases: Simon as an instructed student, and Emil as a non-instructed student. On the one hand, the individual modelling processes are illustrated in a diagram, where the process and the duration of the phases can be seen (Figure 9). The white bars in the diagram show that at this point the individual is not in any phase, i.e., does not make any observable action. This representation provides an overview in comparison to a modelling cycle. It shows which phases follow one another and how long the individual was in the phase. The relative number of occurrences of each phase is shown (Figure 10).



Individual modelling route of Simon.



Individual modelling route of Emil.

Figure 9. Individual modelling routes in a diagram (the meaning of the colours can be found in Figure 10)

Emil does not take much time at the beginning to understand the problem of the modelling task and does not deal with the problem intensively. Simon, on the other hand, deals with the problem for a long time at the beginning and turns back to the phase *real situation* again (Figure 9).

The comparison of Simon and Emil illustrates the result of the statistical test: instructed students thus stay significantly longer and turns more often in the phase *real situation*, since it was made clear in the instruction that understanding the task is the basis for the solution process.

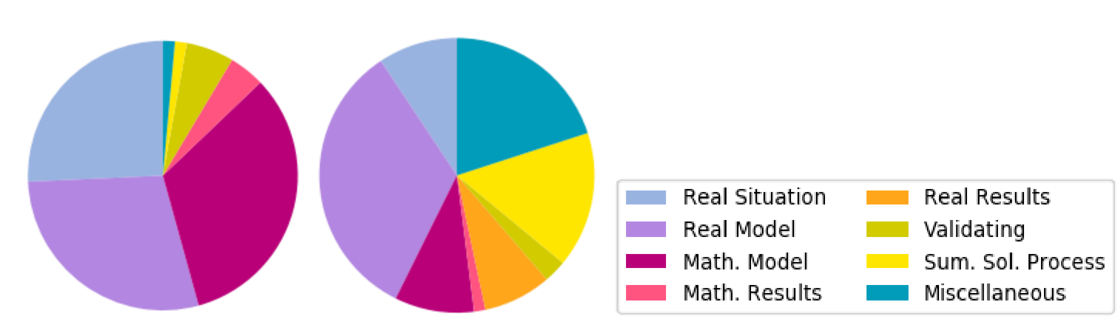


Figure 10. Relative number of the occurrences of the phases of Simon (left) und Emil (right)

The *mathematical world* – consisting of the phases *mathematical model* and *mathematical solution* – has a low part in individual modelling routes without instruction compared to those with instruction, seen over the entire modelling process. A look at the *mathematical world* in Emil's modelling process shows that the relative number and the relative duration of the phases *mathematical model* and *mathematical solution* are low. The modelling route in Figure 10 underlines that Emil did not spend much time in these phases. In contrast, the part of *mathematical world* is higher for Simon: the relative part and the relative duration of *mathematical model* and *mathematical solution* are higher than for Emil. This can be attributed to the influence of the given instruction about ideal-typical modelling processes. Within the instruction, the creation of the *mathematical model* was first discussed using an example and then generalised using the modelling cycle. The model becomes more complex the more assumptions are made, i.e., the more properties of reality are included in the model, the more difficult it becomes. The instruction made the students aware that the *mathematical model* is more complex than it appears at first sight. This has an effect on the relative number and the relative duration in the modelling process that the students engage more intensively with mathematics.

A new feature of the modelling process that has not yet been investigated are the phases *summary solution process* and *miscellaneous*. The relative number and relative duration of *miscellaneous*, i.e., the interactions that are context-independent, are higher for Emil than for Simon. Since the relative number as well as the relative duration are higher in each case, Emil is busy not talking about the task often and for a long time. Looking at the individual modelling process in Figure 9, it becomes apparent that he is often preoccupied with things other than the modelling task towards the end (from minute 16). It can be concluded from this that he lacked knowledge about the modelling process and was therefore unable to proceed in a structured manner. In contrast to Emil, Simon's behaviour is different: the relative number as well as the relative duration are low. This means that he rarely deviates from the processing in the solution process, and if he deals with something else, then this is brief and he switches back to a phase of the modelling cycle.

Not only the context-independent interactions, but also the phase *summary of the solution process* is higher in the relative number and relative duration for non-instructed students than for instructed students. Especially towards the end of the solution process, Emil is frequently busy with the recapitulation of the solution process which takes a long time in total. In contrast, Simon is mostly in the *mathematical model*. It can be concluded from this that non-instructed students limit themselves to their solution, while instructed students continue to be interested in an (improved) solution.

Based on the comparison of Simon’s and Emil’s modelling routes and the results of the statistical test, it can be seen that instructed students deviate less from the processing of the modelling task than the non-instructed students. This means that the given knowledge about the modelling process has an influence on the solution process, as students work on the task in a more structured way.

Based on the structure of the modelling process, it could be determined that the instructed students were oriented towards the solution plan. Based on the results of the statistical test, we can generalise the structure of the individual modelling routes of the instructed and non-instructed students:

- Individuals with an instruction about ideal-typical modelling processes take time to understand the task. They engage often and for a long time in the *mathematical world* by formulating a *mathematical model* and solving it. They engage in the solution process and rarely recapitulate their previous solution. Possibly because they know the modelling process, they rarely and only briefly digress from the topic of the task.
- Individuals without an instruction about ideal-typical modelling processes create a model after a short orientation phase. This leads to problems in creating and solving the *mathematical model*, which is why they only deal with the mathematics for a short time. Due to the lack of an instruction, the individuals often talk for a long time about topics that have nothing to do with the task. Possibly because they are unable to proceed in a structured manner, they summarise their already elaborated solution to the task instead of further improving their model.

Further Perspectives

In this article, we point out the differences in the structure of the modelling routes of instructed and non-instructed students. They differed significantly in the relative number and relative duration of the phases *real situation*, *mathematical model*, *mathematical solution*, *summary solution process* and *miscellaneous*. We interpret these differences as due to the influence of the given instruction. Since we have so far looked at the relative number and relative duration of the phases aggregated over the entire modelling process, we intend to examine in a next step at which points in time certain phases occur more frequently and more often. In order to be able to investigate the structure of individual modelling routes even more precisely, we will develop a concept to describe them on the basis of patterns.

Acknowledgments

This study was supported by the European Social Fund (ESF) project SchuMaMoMint.

References

Ärlebäck, J. B., & Albarracin, L. (2019). An extension of the MAD framework and its possible implication for research. *Eleventh Congress of the European Society for Research in Mathematics Education.* Utrecht, Netherlands. https://hal.archives-ouvertes.fr/hal-02408679/document

Beckschulte, C. (2019). *Mathematisches modellieren mit lösungsplan: Eine empirische untersuchung zur entwicklung von modellierungskompetenzen*. Springer. https://doi.org/10.1007/978-3-658-27832-8

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological), 57*(1), 289-300. https://www.jstor.org/stable/i316032

Blum, W. (1996). Anwendungsbezüge im mathematikunterricht – Trends und perspektiven. In G. Kadunz, H. Kautschitsch, G. Ossimitz, & E. Schneider (Eds.), *Trends und perspektiven* (pp. 15-38). Hölder-Pichler-Tempsky.

Blum, W. (2007). Mathematisches modellieren – zu schwer für schüler und lehrer?. In *Beiträge zum mathematikunterricht* (pp. 3-12). WTM.

Blum, W. (2015). Quality teaching of mathematical modelling: What do we know, what can we do? In S. J. Cho (Ed.), *The Proceedings of the 12th International Congress on Mathematical Education - Intellectual and attitudinal challenges* (pp. 73-96). Springer. https://doi.org/10.1007/978-3-319-12688-3_9

Blum, W., & Borromeo Ferri, R. (2009). Mathematical modelling: Can it be taught and learnt? *Journal of Mathematical Modelling and Application*, *1*(1), 45-58.

Blum, W., & Leiß, D. (2006). "Filling up" – The problem of independence-preserving teacher interventions in lessons with demanding modelling tasks. In M. Bosch (Ed.), *CERME-4 – Proceedings of the Fourth Congress of the European Society for Research in Mathematics Education* (pp. 1623-1633). Sant Feliu de Guíxols, Spain.

Borromeo Ferri, R. (2006). Theoretical and empirical differentiations of phases in the modelling process. *ZDM – The International Journal for Mathematics Education*, *38*(2), 86–95. https://doi.org/10.1007/BF02655883

Borromeo Ferri, R. (2007). Modelling problems from a cognitive perspective. In C. Haines, P. Galbraith, W. Blum, & S. Khan (Eds.), *Mathematical modelling: Education, engineering and economics. ICTMA12* (pp. 260-270). Horwood Publishing.

Borromeo Ferri, R. (2010). On the influence of mathematical thinking styles on learners' modeling behavior. *Journal für Mathematik-Didaktik, 31*(1), 99-118. https://doi.org/10.1007/s13138-010-0009-8

Borromeo Ferri, R. (2011). *Wege zur Innenwelt des mathematischen modellierens. Kognitive analysen zu modellierungsprozessen im mathematikunterricht.* Vieweg+Teubner.

Borromeo Ferri, R. (2018). *Learning how to teach mathematical modeling – in school and teacher education*. Springer. https://doi.org/10.1007/978-3-319-68072-9

Brand, S., & Vorhölter, K. (2018). Holistische und atomistische vorgehensweisen zum erwerb von modellierungskompetenzen im mathematikunterricht. In S. Schukajlow & W. Blum (Eds.), *Evaluierte lernumgebungen zum modellieren* (pp. 119-142). Springer.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-46.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Lawrence Erlbaum Associates.

Döring, N., & Bortz, J. (2016). *Forschungsmethoden und evaluation. In den sozial- und humanwissenschaften.* Springer. https://doi.org/10.1007/978-3-642-41089-5

Flavell, H. J. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychologist, 34*(10), 906-911.

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations and interpretation. *Journal of Experimental Psychology: General, 141*(1), 2-18. https://doi.org/10.1037/a0024338

Galbraith, P., & Stillman, G. (2006). A framework for identifying student blockages during transitions in the modelling process. *Zentralblatt für Didaktik der Mathematik, 38*(2), 143-162. https://doi.org/10.1007/BF02655886

Greefrath, G. (2014). Lösungshilfen für modellierungsaufgaben. In I. Bausch, G. Pinkernell, & O. Schmitt (Eds.), *Unterrichtsentwicklung und kompetenzentwicklung – Festschrift für regina bruder* (pp. 131-140). WTM.

Greefrath, G. (2018). *Anwendungen und modellieren im mathematikunterricht. Didaktische perspektiven zum sachrechnen in der sekundarstufe*. Springer. https://doi.org/10.1007/978-3-662-57680-9

Greefrath, G., Kaiser, G., Blum, W., & Borromeo Ferri, R. (2013). Mathematisches modellieren – Eine einführung in theoretische und didaktische hintergründe. In R. Borromeo Ferri, G. Greefrath, & G. Kaiser (Eds.), *Mathematisches modellieren für schule und hochschule. Theoretische und* *didaktische Hintergründe* (pp. 11-37). Springer. https://doi.org/10.1007/978-3-658-01580-0_1

Kaiser, G., Blum, W., Borromeo Ferri, R. & Greefrath, G. (2015). Anwendungen und modellieren. In R. Bruder et al. (Eds.), *Handbuch der mathematikdidaktik* (pp. 357-383). Springer. https://doi.org/10.1007/978-3-642-35119-8_13

Kaiser, G., & Stender, P. (2013). Complex modelling problems in co-operative, self-directed learning environments. In G. Stillman, G. Kaiser, W. Blum, & J. Brown (Eds.), *Teaching mathematical modelling: Connecting to research and practice* (pp. 277–293). Springer. https://doi.org/10.1007/978-94-007-6540-5_23

Kolbe, M., & Boos, M. (2018). Observing group interaction: The benefits of taking group dynamics seriously. In E. Brauner, M. Boos, & M. Kolbe (Eds.), *The Cambridge handbook of group interaction analysis* (pp. 68–85). Cambridge University Press. https://doi.org/10.1017/9781316286302.005

Leuders, T. (2007). Wenn es mathematikern zu bunt wird: Färbeprobleme. In S. Hußmann & B. Lutz Westphal (Eds.), *Kombinatorische optimierung erleben. In schule und unterricht* (pp. 131-170). Vieweg.

Maaß, K. (2004). *Mathematisches modellieren im unterricht. Ergebnisse einer empirischen studie.* Franzbecker.

Matos, J. F., & Carreira, S. (1997). The quest for meaning in students' mathematical modelling activities. In S. K. Houston, W. Blum, I. Huntley, & N. T. Neill (Eds.), *Teaching and learning in mathematical modelling* (pp. 63-75). Albion.

Niss, M., Blum, W., & Galbraith, P. (2007). Introduction. In W. Blum, P. Galbraith, H-W. Henn & M. Niss (Eds.), *Modelling and applications in mathematics education. The 14th ICMI Study* (pp. 3-32). Springer.

Pollak, H. (1979). The interaction between mathematics and other school subjects. In UNESCO (Eds.), *New trends in mathematics teaching IV* (pp. 232-248). UNESCO.

Ruzika, S., & Schneider, L. (2020). Modellierungsprozesse erfassen, darstellen und analysieren. In H-S. Siller, W. Weigel, & J. F. Wörler (Eds.), *Beiträge zum mathematikunterricht 2020* (pp. 1201-1204). WTM.

Schukajlow, S., Kolter, J., & Blum, W. (2015). Scaffolding mathematical modelling with a solution plan. *Zentralblatt für Didaktik der Mathematik, 47*, 1241-1254. https://doi.org/10.1007/s11858-015-0707-2

Stillman, G. (2004). Strategies employed by upper secondary students for overcoming or exploiting conditions affecting accessibility of applications tasks. *Mathematics Education Research Journal, 16*(1), 41-71. https://doi.org/10.1007/BF03217390

Stillman, G. (2011). Applying metacognitive knowledge and strategies in applications and modelling tasks at secondary school. In G. Kaiser, W. Blum, R. Borromeo Ferri, & G. Stillman (Eds.), *Trends in teaching and learning of mathematical modelling* (pp. 165-180). Springer. https://doi.org/10.1007/978-94-007-0910-2_18

Stillman, G., & Galbraith, P. (1998). Applying mathematics with real world connections: metacognitive characteristics of secondary students. *Educational Studies in Mathematics, 36*, 157-189. https://doi.org/10.1023/A:1003246329257

Strauß, A., & Corbin, J. (1996). *Grounded theory. Grundlagen qualitativer sozialforschung*. Beltz.

Vorhölter, K., & Kaiser, G. (2016). Theoretical and pedagogical considerations in promoting students' metacognitive modeling competencies. In C. R. Hirsch (Ed.), *Mathematical modeling and modeling mathematics* (pp. 273-280). National Council of Teachers of Mathematics.

Vorhölter, K., Krüger, A., & Wendt, L. (2019). Metacognition in mathematical modelling – An overview. In S. A. Chamberlin, & B. Sriraman (Eds.), *Affect in mathematical modeling* (pp. 29-51). Springer. https://doi.org/10.1007/978-3-030-37673-4_27

Zöttl, L., Ufer, S., & Reiss, K. (2011). Assessing modelling competencies using a multidimensional IRT approach. In G. Kaiser, W. Blum, R. Borromeo Ferri, & G. Stillman (Eds.), *Trends in teaching and learning of mathematical modelling* (pp. 427-437). Springer. https://doi.org/10.1007/978-94-007-0910-2_42