ORIGINALARBEIT

Modelling with Heuristic Worked Examples in the KOMMA Learning Environment

Luzia Zöttl · Stefan Ufer · Kristina Reiss

Abstract Modelling competency is considered to be an important part of mathematical competency. Although much work has been done with respect to the development and dissemination of modelling tasks, there is hardly any empirical evidence how these tasks should be integrated in the mathematics classroom. In this article we present a teaching approach based on heuristic worked examples which turned out to be a promising way of supporting initial skill acquisition within the field of modelling. We will provide an overview how the learning environment (KOMMA) was conceptualized and will present first results of a large-scale field study examining the effectiveness of example based learning for initial skill acquisition in the field of modelling. The results presented here take into account short-term as well as long-term effects on the learning of 8th-grade students. They suggest, that the participants' modelling competencies increased significantly during the training but that long-term effects were much smaller.

L. Zöttl
Mathematisches Institut, Ludwig-Maximilians-Universität München, Theresienstraße 39, 80333 München, Deutschland

S. Ufer (✉) · K. Reiss
TUM School of Education, Technische Universität München, Schellingstraße 33, 80799 München, Deutschland
e-mail: ufer@math.lmu.de

Modellieren mit heuristischen Lösungsbeispielen in der Lernumgebung KOMMA

Zusammenfassung Modellierungskompetenz wird als wichtiger Teil mathematischer Kompetenz betrachtet. Trotz umfangreicher Forschungen zur Entwicklung und Dissemination von Modellierungsaufgaben gibt es kaum empirische Ergebnisse dazu, wie diese Aufgaben in den Mathematikunterricht integriert werden sollen. Wir beschreiben einen auf heuristischen Lösungsbeispielen basierenden instruktionalen Ansatz, der sich als vielversprechend für den anfänglichen Fähigkeitserwerb beim Modellieren erwiesen hat. Wir beschreiben das Konzept der verwendeten Lernumgebung (KOMMA) und berichten erste Ergebnisse einer größeren Feldstudie zur Untersuchung der Effektivität von beispielbasierten Lernumgebungen für den anfänglichen Erwerb von Modellierungskompetenzen. Die Ergebnisse beschreiben kurzfristige und langfristige Lernzuwächse bei Schülerinnen und Schülern der achten Jahrgangsstufe. Sie legen nahe, dass die Modellierungskompetenzen der Schülerinnen und Schüler während der Intervention signifikant zunahmen, langfristige Effekte aber deutlich kleiner ausfielen.

1 Introduction

Mathematical modelling is regarded an important topic in the mathematics classroom. It encompasses aspects like as constructing an adequate model for a specific problem, applying mathematical knowledge in this situation, and interpreting it properly. Its prominent role is emphasized by the fact that modelling is an important part of the PISA mathematics tests (OECD 2003) and one of six general mathematical competencies that have been identified in the German standards for secondary school mathematics (Kultusministerkonferenz 2003).

These standards for school mathematics provide numerous examples for what is meant by modelling in the mathematics classroom, however, there are hardly any empirically validated ideas how the acquisition of modelling competency can be supported. Ways of acquiring this competency were addressed in the research study KOMMA (Kompendium Mathematik).[1] In particular, this study aimed at developing and evaluating a computer learning environment for 8th-graders with a competency-oriented implementation of mathematical content (Reiss et al. 2007). As 8th-grade students usually only have minor experience in modelling, KOMMA was designed to allow initial skill acquisition in this field. In order to introduce heuristic strategies necessary for a successful modelling process, the learning environment was based on heuristic worked examples (cf. Reiss and Renkl 2002). These examples had already turned out to be appropriate for initial skill acquisition in the field of mathematical proof and argumentation (see Sect. 2.1).

In this paper we will describe the principles of heuristic worked examples within the field of modelling and give an idea how this teaching approach was implemented in the KOMMA learning environment. Moreover, we will present first results of a

---

large-scale field study on the effectiveness of the example-based learning. We will concentrate on geometry learning in a computer-assisted environment.

2 Worked Examples

2.1 The Principle of Worked Examples

Worked examples have been investigated intensively in the last few years as they are regarded to provide students with a perfect solution of a problem. They typically comprise a problem statement, a step-by-step procedure for solving the problem, and the solution itself and have been proven to be effective for initial skill acquisition in well-structured domains (Atkinson et al. 2000; for an overview, see e.g. Sweller et al. 1998). These worked examples provide an expert's problem-solving model and implicitly show how other similar and isomorphic problems might be solved by using analogical transfer. They provide a solution of a problem by presenting a straight forward problem-solving process. A problem solver is asked to solve similar problems by imitating the solution steps presented. It is usually sufficient for problem solving to manipulate the data in the initial worked example. This type of analogical transfer is called transformational analogy (Carbonell 1986) because only a simple modification of a typical example is required.

The effectiveness of worked examples is explained by Sweller et al. (1998) within Cognitive Load Theory. According to this theoretical approach, a major part of the cognitive resources in problem-solving situations is used for finding the right solution steps. Problems presented as worked examples already have a solution, and therefore, lower cognitive capacities are needed and can be used to better understand the presented problem solution and to construct adequate mental structures (Sweller 2003; Paas et al. 2003).

This theoretical explanation is supported by findings that students perform well on similarly structured tasks after studying algorithmic worked examples. Several studies showed that learners who worked with examples outperformed those who had to solve identical problems on their own (see e.g. Sweller et al. 1998, pp. 273–275). In addition, learning with worked examples is not only more effective with respect to the learning outcome, but also with respect to the time spent for learning. Studies show that students learning with worked examples achieved similar or even better results and in addition needed less working time compared to learners who solved the same number of problems on their own (Sweller and Cooper 1985; Zhu and Simon 1987; Carroll 1994). An even more important advantage of this method is the popularity of examples compared to abstract rules or instructions. The results of several studies verify that learners strongly prefer examples (LeFevre and Dixon 1986; Recker and Pirolli 1995).

It should be mentioned that positive effects were primarily identified for novice learners. Moreover, the solution of a problem had to be presented in adequate detail and had to take into account the individual prior knowledge of the learner. Accordingly, the studies suggest that learners need more guidance for initial skill acquisition than for the extension an already existing expertise. Furthermore, the use of worked

examples may even induce negative consequences for advanced learners. This *expertise reversal effect* is shown in several investigations (e.g., Sweller 2003). The effect emerges presumably from the situation that the processing of unnecessary or redundant information consumes cognitive capacity, too. According to the Cognitive Load Theory (Sweller et al. 1998), this interferes negatively with effective learning.

2.2 Heuristic Worked Examples

Most research on worked examples was performed in well-structured domains and on relatively simple tasks. However, much learning takes place in less-structured or ill-structured domains. Students who are presented ideal solutions in such domains may not be able to grasp the relevant ideas of the problem solution. Accordingly, Reiss and Renkl (2002) suggested *heuristic worked examples* which do not only provide a solution but take into account the solution steps. In less-structured or ill-structured domains, it is not sufficient for learning to emulate a presented algorithm. As problems will hardly be solved by using specific algorithms, the development of an individual solution procedure for each problem is required. However, in such cases, memorized or given examples may help to solve new problems as well. An approach to explain how examples in such fields might support learners in problem-solving situations is the derivational analogy (Carbonell 1986; Schelhorn et al. 2007). This kind of analogical transfer refers to an adaptation of a memorized or given problem-solving procedure and by using the structure of the procedure's subgoals. These guidelines should help finding chains of reasoning, decision sequences or successive solution steps. Thus, even in less-structured domains working with suitable examples might help to structure the problem-solving process and find relevant strategies and heuristics for similar problems.

Reiss and Renkl (2002) developed the principle of heuristic worked examples for the domain of mathematical proof and argumentation. Their examples involve a domain-specific process model which helps students structure their problem-solving process. This process model is based on an expert model of mathematical proof (Boero 1999) which was adapted to be more appropriate for students at the lower secondary level. It describes the different steps and corresponding heuristic strategies of an expert working on a mathematical proof.

As learning from heuristic worked examples should introduce students to problem solving in a certain domain, Reiss and Renkl (2002) used a realistic instead of an optimized output-oriented problem-solving process in their heuristic worked examples. In order to present a realistic problem-solving procedure, heuristic worked examples include tentative and explorative steps and explain heuristics and heuristic tools. Accordingly, they can be regarded as process-oriented instead of product-oriented (see also van Gog et al. 2004, and their concept of process-oriented worked examples).

There is empirical evidence that heuristic worked examples enhance students' understanding of mathematics. In particular, heuristic worked examples have been shown to be effective for problems requiring mathematical argumentation and proof. Field studies with 8th-grade high school students (Reiss et al. 2006) and first-year university students (Hilbert et al. 2008) revealed that working with these examples fostered the development of proof competency.

Fig. 1 Modelling cycle adapted
from Blum and Leiss (2006)
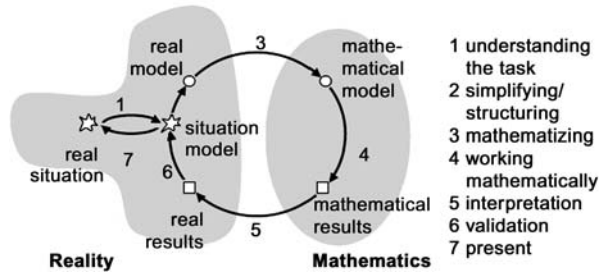
3 Modelling and Modelling Competency

3.1 The Modelling Process

What research means by mathematical modelling varies considerably between different groups (see Kaiser and Sriraman 2006, for an overview). The differences are mainly due to the different goals researchers emphasize with respect to students' modelling activities. However, there is consensus that working on real world problem is important and that students have to move between reality and mathematics. There are different points of view concerning the level of authenticity and complexity a modelling task should offer in order to be regarded a modelling task. The KOMMA project is based on a modelling perspective which was elaborated by Blum (1996). From this point of view, the most important feature of a modelling task is not its level of authenticity and complexity but its relevance for the students. The modelling process can be seen as a sequence of seven phases which are summarized in an idealized cycle (see Fig. 1, Blum and Leiss 2007).

The modelling process starts with a real-world problem. First, it is essential to understand a problem task in order to build an *idiosyncratic* situation model. After simplifying and structuring this model, the solver attains a so-called *real model* of the problem situation. It becomes a *mathematical model* by mathematizing, which means translating it into mathematics. The aim of the subsequent process steps is to solve the resulting mathematical problem. This step is called *working mathematically*. The interpretation of this mathematical solution corresponds to its re-translation in the real context. Results have to be checked by a validation step. It should answer the question if the primary problem task is solved in a satisfactory manner. If it does so, the solution of the problem needs to be presented in an appropriate way. This is considered as being the last step of the modelling process. If the results do not fit into the real context, the problem solver has to repeat the modelling process or parts of it in order to get or optimize the solution (cf. Leiss 2007).

3.2 Modelling Competency

The development of a learning environment for enhancing students' modelling competency as well as the construction of an adequate test instrument should be based on an appropriate definition of the term *modelling competency*. This definition should take into account the psychological perspective for example by accepting the perspective on competency suggested by Weinert (2001). Moreover, it should take into

account the mathematics education perspective by including specifics of the modelling process.

Seeing both perspectives, modelling competency can be defined as the ability and readiness to solve an appropriate problem. However, this definition can be further elaborated by looking at the modelling cycle and by educing corresponding subcompetencies from its different phases. As a consequence, a component model should result which focuses on the different components of modelling competency. According to Blum and Kaiser (1997) the subcompetencies encompass understanding of the real problem, setting up of a model based on reality, excerpting a mathematical model from the real model, answering mathematical questions within this mathematical model, interpreting mathematical results in a real situation and validating the solution.

Research suggests that these subcompetencies are important prerequisites for modelling competency but further aspects should be considered, e.g. the coordination of those subcompetencies (Treilibs et al. 1980). As suggested by Reiss and Renkl (2002), strategic knowledge is required for the process of mathematical proof as well as for successful modelling. This aspect is included in a further model of competency by Jensen (2007). The model identifies the following three aspects of modelling competency.

> *Degree of coverage*, indicating which *aspects* of the competency someone can activate and the degree of *autonomy* with which this activation takes place.
> *Radius of action*, indicating the spectrum of *contexts and situations* in which someone can activate the competency.
> *Technical level*, indicating how *conceptually and technically advanced* the mathematics is that someone can integrate relevantly in activating the competency. (Jensen 2007, pp. 143–144)

According to this model, a person with high modelling competency should be able to solve modelling tasks in different contexts and situations, using conceptually and technically advanced mathematical entities and tools. Thereby, solving a modelling task means running autonomously, i.e. without being prompted to do so, through the various steps in the modelling process.

The theoretical thoughts presented here may not only influence the construction of a learning environment but the construction of test instruments as well. Accordingly, in order to assess students modelling competency, test items should include varying degrees of coverage. In particular a useful test should include items which cover the whole modelling process as well as items focusing only on parts of this process or prompting the sequence of relevant phases. On the one hand, this guarantees to measure modelling competency of low-achieving students, too, who are not able to autonomously perform a complete modelling task. On the other hand, test results thus provide detailed information about students' specific strengths and weaknesses concerning specific subcompetencies. The other two aspects (radius of action and technical level) require that test items should differ in context and the mathematical tools needed for their solution.

3.3 Teaching Approaches to Modelling

In the field of modelling, there are different teaching and learning approaches. An important technique may be described as the "*holistic approach in which students learn*

through experiences of complete case studies in mathematical modelling" (Haines et al. 2003, p. 42). At the beginning, simple modelling tasks should be used which can be solved directly. With increasing competency, students should work on more difficult situations. Moreover, there is the *atomistic* approach which aims at "*developing students' mathematical modelling competency* [...*by concentrating*] *on the processes of mathematizing and analysing models mathematically*." (Blomhøj and Jensen 2003, p. 128). This approach focuses on the process steps *mathematizing*, *working mathematically* and *interpreting* and thus covers only part of the modelling competency (Blomhøj and Jensen 2003, p. 130). A serious problem of the atomistic approach is, however, that only certain subprocesses are taught. Thus, a rigid atomistic approach, even if it is time-saving, does not allow the development of a high degree of coverage. In order to avoid these negative consequences, Blomhøj and Kjeldsen (2006) proposed a balance of the two approaches.

A third teaching approach proposes the presentation of exemplary models. "*It often involves the presentation, discussion and analysis of several applications in particular fields addressing a basic applied mathematics problem* [...]" (Haines and Crouch 2006, p. 1656). Modelling courses following this approach are usually based on a teacher-centered presentation of several real situations concerning the same underlying mathematical model. Whereas the latter approach may be seen as classical, there is another less known approach including examples. Legé (2005) performed a case study with two experimental groups from two comprehensive high schools. The first group worked on different exemplary solutions of a modelling task (planning vacations, given a limited time slot and limited money). The students were presented mathematical models aiming at different aspects of the task and were asked to answer questions concerning the reproduction of calculations, the explanation of assumptions underlying the models, and the comparison of the different models. They were prompted to self-explain certain aspects of the presented solutions (for further information see Sect. 4.1). The students of the second group had to model the same task without exemplary solutions. In order to assess the students' learning success, a second modelling problem was used, in which all students had to construct their own models. However, Legé (2005) could not identify important differences between the groups. He summarizes that both approaches were feasible but a general superiority of one approach could not be detected. It is an open question whether this result can be replicated with a larger representative sample and other contexts and classes.

4 The KOMMA Learning Environment

The research about mathematical modelling described in Sect. 3 is largely based on theory. Empirical findings on effects of modelling instruction in mathematics classrooms are rare. Moreover, results are hardly based on quantitative data but are mostly restricted to the description of small qualitative case studies. This was the starting point for KOMMA, a project aiming at the implementation and evaluation of a learning environment for mathematical modelling. This learning environment was based on the idea of heuristic worked examples and took advantage of a definition of competency which regarded cognitive as well as affective components.

Fig. 2 Exemplary modelling task in the KOMMA learning environment

The KOMMA learning environment was supposed to enhance students' understanding for geometry and statistics in grade 8. We chose these content areas because geometry has been intensively taught since the first years of school whereas statistics is a relatively new subject in grade 8. As a second aspect of differentiation, we implemented KOMMA in a paper and pencil as well as in a computer-based version. Finally, we included levels of self-regulation, namely a higher one (students were able to choose problems according to their interest and competency) and a lower one (students had to work in a fixed sequence of problems, however, every problem had to be solved self-regulated).

We will describe in this paragraph important aspects of the implementation of the KOMMA learning environment for geometry as a computer-based instrument. Moreover, we will concentrate on the implementation with a low level of self-regulation and a higher guidance of the program. The learning environment aimed at introducing measurement of area and circumference of the circle and presenting adequate applications and modelling tasks.

4.1 Examples and Tasks

Heuristic worked-out examples: The KOMMA geometry learning environment included four heuristic worked-out examples. They were presented as dialogues of two fictitious persons solving a modelling task (see Fig. 2 for an exemplary modelling task).

We used an adapted version of the modelling cycle (see Sect. 3.1) as a process model in order to structure all worked examples in a similar way. The model was reduced to three subprocesses: (1) *Understanding the Task*, which included the first
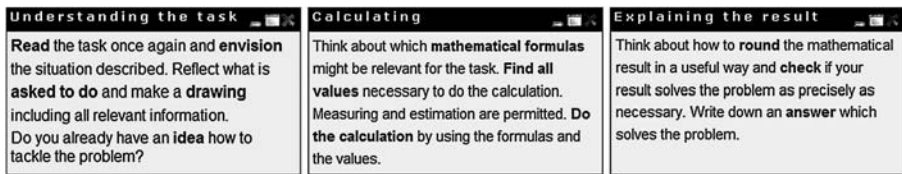
Understanding the task	Calculating	Explaining the result
Read the task once again and **envision** the situation described. Reflect what is **asked to do** and make a **drawing** including all relevant information. Do you already have an **idea** how to tackle the problem?	Think about which **mathematical formulas** might be relevant for the task. **Find all values** necessary to do the calculation. Measuring and estimation are permitted. **Do the calculation** by using the formulas and the values.	Think about how to **round** the mathematical result in a useful way and **check** if your result solves the problem as precisely as necessary. Write down an **answer** which solves the problem.

Fig. 3 Process model including helpful strategies

three phases of the modelling cycle, (2) *Calculating*, which concerned the fourth phase, and (3) *Explaining the Result*, which comprised the last phases (Zöttl and Reiss 2008). This process model was used to structure the solution presented within the worked example. Every process step was presented on a separate page starting at the top with a short general description of the step calling the attention to helpful strategies as e.g. making a drawing of the situation (see Fig. 3).

Within the dialogue, the two fictitious problem solvers explained their ideas, heuristic strategies, and heuristic tools throughout the whole modelling process (see Fig. 4 for an illustration). As heuristic examples usually try to integrate explanations on the level of students' understanding as well as on an expert's level, we implemented one of the two persons discussing the modelling task as a novice (here: Tobias), the other person however as a more advanced learner (here: Kristina).

The effectiveness of learning with worked examples depends on the level of self-explanation activities of the learners during their work. Therefore, we implemented *self-explanation prompts* (titled: "working instruction") in order to support students who would not start self-explaining on their own while reading a worked example (cf. Chi et al. 1989). Prompts called on learners to self-explain a specific aspect of the given solution at a certain point of the worked example. The implementation took into account that examples might be read superficially and without deeper understanding (Chi et al. 1989). However, successful learning with worked examples can only be expected if a person uses his or her cognitive capacity for the concentration on important aspects of the problem solution.

In the literature, different types of prompts are discussed (see Renkl 2002a, for an overview). We integrated primarily the *anticipative prompts* (the first instruction concerns the first process step, the second one concerns the third process step). They ask the learner to predict the next solution step in the example studied presently. The effectiveness of this type of prompts was shown by Stark (1999) for probability calculation. He presented partly incomplete worked examples and, as a feedback after completing the gaps, the whole solution was presented. This type of prompts was integrated in all subprocesses of our worked examples. They were chosen in order to ensure a deeper processing of the presented modelling process.

The examples also included some *principle-based prompts* in the third subprocess which were supposed to support students in developing knowledge about the final steps of the modelling process (interpretation and validation). These prompts aimed at reflecting on the necessity to do these steps. In studies by Atkinson et al. (2003) as well as by Aleven and Koedinger (2002) these principle-based prompts already showed to be effective with well-structured material.

Fig. 4 Detail of the
problem-solving dialogue

4.1.1 Exercise Examples

We implemented exercise examples in the learning environment to facilitate the transition from working with complete worked examples to independent problem solving and to avoid an expertise reversal effect (see Sect. 2.2). In these exercises, the complete modelling process with all its solution steps is hidden in the beginning. However, it can be retrieved step by step on demand via the *help* button (see Fig. 5).

Thus, the examples were kind of incomplete worked examples (Sweller et al. 1998) which should encourage students to decide how much guidance of the underlying worked example they needed and thus provided the opportunity to be used in a self-adaptive way.

Renkl (2002b) summarizes so-called *SEASITE*[2] *principles* for instructional explanations in example-based computer-learning environments: "As much self-explanation as possible, as much instructional explanation as necessary. [...] Provide feedback. [...] Provision on learner demand. [...] Minimalism. [...] Progressive help. [...] Focus on principles [...]". According to these principles, feedback in the exercise examples was structured as a progressive help system (see Fig. 6). Thus, a student could check the solution for a specific step in the modelling process (retrievable via the exclamation mark) or could get more information about the step and the associated problem-solving process if required (retrievable via the question mark). The feedback progressively increased in its details, meaning that the complete solution of the modelling task could be faded in step by step if the learner wanted to do so. So a student found a very detailed explication of the solution on the last and

---

[2]SEASITE means self-explanation activity supported by instructional explanations.

Task: Mosaicked table

Mathew and Andrea would like to refurbish a small garden table. They'd like to glew special tesseras onto the table, as they have seen it in a do-it-yourself program. In the Internet, suitable tesseras were offered. Tesseras are quite expensive. Therefore they try to find out the exact number they have to buy.

Shopping List and Tool:

- mosaicked table from the do-it-yourself program ⓘ
- Tool to measure the tabletop. (Here you get information how to use the program MEASURING. ⓘ)
  ✋ MEASURING
- Green tesseras from the Internet ⓘ

Here you can check your result: ⓠ
Here you get help for solving the problem: [I Ielp]

Fig. 5 Exemplary exercise example in the KOMMA learning environment

Help with step 1:   [Understanding the task]

Help with step 2:

Think about what **mathematical formulas** might be relevant for the task. **Find all values** necessary to do the calculation. Measuring and estimation are permitted. **Do the calculation** using the formulas and the values.

1. step: Estimate the **size of the tabletop!** ⓠ ❷
2. step: Estimate the **area which can can be covered with one package of tesseras!** ⓠ ❷
3. step: Find out **the needed number of packages!** ⓠ ❷

Help with step 3:   [Explaining the result]

Fig. 6 Progressive help within the exercise examples

most elaborated level of the help system. Accordingly, the example-based learning environment was designed to be appropriate for low-achieving as well as for high-achieving students, as the feedback was provided only on learners demand and meant as a progressive help system permitting support for several aptitude levels.

The extensive feedback possibilities may result in a decreasing willingness to work hard on the problems as a correct solution could easily be retrieved and read without

effort (Renkl 2002b, pp. 534–535). However, the help system avoided difficulties in the learning process. This is important since during the geometry course students worked on their own without additional external support. In particular, the teachers were asked not to intervene whenever it was possible to do so.

4.1.2 Technical Tasks

As modelling is a complex task, we tried to avoid cognitive overload when working with heuristic worked examples. Therefore, we proposed the acquisition respectively recapitulation of the required algorithmic skills (i.e. computing circular area). Pollock et al. (2002) propose such a separation of subcomponents with very complex material in well-structured domains. This separation seemed reasonable, too, with respect to the different learning approaches in the field of modelling (see Sect. 3.3), as an appropriate balance of the holistic approach (implemented in worked and exercise examples), and the atomistic approach (implemented by the separation of the technical aspects) was considered to be ideal. Thus, during the first unit, the students worked on algorithmic examples and mathematical tasks concerning the circular area.

4.1.3 Additional Support Features

As the students worked autonomously throughout the geometry course there were several features supporting their work. For instance, self-tests consisting of several items were integrated in the course and presented at the end of each unit. These tests should help students to evaluate their learning progress. Additionally, a learning diary and a schedule were offered in order to document the learning processes. Both features were meant to initiate a reflection of the individual learning on a metacognitive level. However, the students were not obliged to use these features, but were reminded that they could use it at the end of each unit.

4.2 Structure of the Geometry Course

The KOMMA learning environment was implemented as a regular classroom instruction in grade 8. It was not intended to introduce the topic but to provide opportunities for exercises and in-depth learning. The geometry course encompassed five teaching units of 45 minutes. The first unit served as an introduction and made the students familiar with the KOMMA software. The students used algorithmic worked examples and tasks concerning the computation of circular area. In order to become acquainted with the modelling cycle and its adapted version (see Sect. 4.1), they were given a short instructional text about modelling and the modelling procedure. After the introductory session, the participants were asked to study heuristic worked examples and corresponding exercise examples for the subsequent four units. Every unit consisted of a worked example and a corresponding exercise example in order to allow a smooth transition into problem solving. Moreover, self-test items were included at the end of the sessions.

Within the second unit the students finished the worked example "information board" (see Fig. 2) and the exercise example "mosaicked table" (see Fig. 5). During

Fig. 7 Spaghetti portioning

the third unit they had to deal with maps. The heuristic worked example encompassed a rough estimation of the area of the almost circular Island of Gran Canaria, whereas the corresponding exercise example asked the students to figure out to how many people could stand in a specific semicircular public place. The subsequent unit was concerned with the computation of the area of a circle when its circumference was known (measured with steps). The area of a lighthouse (worked example) and a donjon (exercise example) had to be estimated taking into account the thick walls as well as a slight conical form. During the last unit, the percentage of wasted power caused by an inadequate cooking pot compared to the size of a hotplate was roughly estimated in the worked example. As an exercise the students had to determine the proportions of a measurement device (see Fig. 7).

Providing sequences of worked examples with varying structure and surface features is considered to be more effective than the use of homogeneous examples (Paas and Van Merriënboer 1994; Quilici and Mayer 1996). Therefore, worked examples and exercise examples differed in context and required mathematical as well as heuristic tools and strategies. The heuristic strategies encompassed, e.g., methods to measure or estimate a relevant, unknown length by using a reference value for a more precise estimation. As a heuristic tool, informative figures were introduced, thus implementing the implications drawn from the different competency models. Furthermore, according to the suggestions given for a holistic modelling approach, we used modelling tasks with an increasing degree of difficulty.

5 Evaluation of the Learning Environment KOMMA

5.1 Research Questions

The main focus of the research project was an investigation of students' learning outcomes within the KOMMA learning environment. In particular, the following research questions were addressed.

Fig. 8 Item "Italian Lake"
(item type 1)

Using a map, Anke would like to estimate the area covered by this Italian lake. Draw onto the map what Anke has to measure to find out the circular area of this lake. **In addition**, write down the formula which is necessary to do the calculation! (You do not have to perform the calculation!)
Attention: Use the same terms in the formula and in the map!

(1) Does learning with worked examples within the KOMMA learning environment enhance the students' modelling competency?

A positive effect was expected because learning with heuristic worked examples already had turned out to be effective in the field geometrical proof. Learning with heuristic examples was assumed to be effective for developing modelling competency as well, since the modelling process and the elaboration of a mathematical proof are characterized by a sequence of certain process steps and thus by a certain heuristic strategy.

(2) Does learning with the worked examples within the learning environment KOMMA enhance long-term modelling competency of students?

Most studies investigating the effectiveness of heuristic or algorithmic worked examples concentrate on short-term effects. Since sustainable learning success is a major aim of education, long-term effects are of specific interest. We expected at least slight long-term effects of the KOMMA learning environment.

5.2 Sample and Method

In this study, KOMMA was implemented as a computer environment for geometry learning. Students were presented worked examples and exercise examples in a specific order. The sample consisted of 316 students of grade 8 from 18 classrooms in nine high-track schools, the German "Gymnasium". There were 171 female and 145 male participants who took part on a voluntary basis. The students joined at least four learning units and participated in all three tests, namely a pretest, a posttest right after the treatment, and a follow-up test about six months later. These tests followed a multimatrix design with two strands (i.e. test version A and B) so that the students worked on different test items at all points of measurement.

The treatment took place during the regular mathematics lessons. The teachers were advised to give no additional mathematics lessons during this period. Moreover, their role was restricted to organisational instructions. In particular, they were asked not to give advice concerning the content of the learning material but to guide their students' work according to the instructions of the program if necessary.

5.3 Test Instrument

The tests encompassed four different types of items. The first type was constructed to measure the subcompetencies needed to run through the first subprocess of the adapted modelling cycle (see Sect. 4.1), i.e. the first three steps of the complete modelling cycle (see Sect. 3.1). Figure 8 shows an exemplary item of type 1. In this task, competencies in understanding (making sense of the text), finding a real model (acknowledging the circular surface), and mathematizing (finding the correct formula) were required to solve these types of items.

Items belonging to type 2 required mathematical knowledge in a narrower sense. They related to the second subprocess (see Sect. 4.1) and thus to the fourth step of the modelling cycle and could be characterized as asking for technical competency. For an exemplary item see Fig. 9.

The subcompetencies needed for the third subprocess (see Sect. 4.1) were measured by items of type 3. They required interpretation of a mathematical result in a specific problem situation and the validation of a problem solution with respect to the underlying model. Thus, they mapped steps 5 and 6 of the modelling cycle. Figure 10 shows an example how items of this type looked like.

Fig. 9 Item "Variation of a square" (item type 2)

Select the correct statement!

"If you triple the side length *a* of a square, the area of the new square is ...

☐ ... three times as big as the area of the original square."

☐ ... nine times as big as the area of the original square."

☐ ... twelve times as big as the area of the original square."

☐ ... not comparable to the area of the original square."

Fig. 10 Item "General Sherman Tree" (item type 3)

In front of the world's biggest Tree (General Sherman Tree) there is an information board shaped like a disk of the tree.

The text on the board is:
Circumference of the trunk (at the base): 31.3 m
Biggest diameter of the trunk (at the base): 11.1 m

Paul wonders: "Given a circumference of 31.3 m, you can compute (using the formula for the circumference of circles) that the diameter is about 10.0 m. How is it possible that the biggest diameter of the trunk is 11.1 m?"

Explain how this is possible!

Fig. 11 Item "Spain" (item type 4)

Portugal has a total area of 92 117.5 sqm.

Estimate the total area of Spain by using the map. Show all your work! (You may draw onto the map.)

Type 4 integrated all aspects and could be regarded as complete modelling tasks. The modelling process with all subprocesses mentioned above was needed for the solution. Figure 11 shows an example for an item of type 4.

The different types of items were equally distributed among all test booklets. Every test contained three items of each type, every modelling test consisted of 12 items. Accordingly, 12 was the maximum raw score to be achieved (one raw score for each task). Time was limited to 30 minutes per test.

As already mentioned, the tests were constructed in a multi-matrix design with two strands (i.e. test version A and B). This design finally led to six different forms. All booklets were linked crosswise. Thus every booklet was linked to each booklet of the other strand by 4 anchoring items.

The different types of items were chosen in order to measure students' modelling competency with respect to diverse levels of degree of coverage. Additionally, the tasks referred to different contexts. Although the learning environment primarily covered the topic *measurement of circular area* from a mathematical point of view it was not limited to this topic. The test instrument was not restricted to this topic, as well. Indeed, the tasks required different mathematical concepts and thus abilities on different technical levels in the field of measurement (e.g., area and circumference of rectangles, triangles and circles). Our test items considered all relevant aspects of modelling competency (see Sect. 3.2).

A common scheme was developed for the coding of open items of each item type and then specialized into a separate coding scheme for each single item. These schemes were designed to provide some additional information about the solution, for example, whether the students stated the answer within the problem context or not, whether they used to correct unit for their answer, and if they failed due to inappropriate estimation of lengths in the problem context (a wide range of acceptable estimations was applied in each case). This additional information will not be part of the analysis in this paper. The detailed coding was used to obtain dichotomous scores for each item. About a third of the tests were coded independently by two persons. The consistency of the dichotomous ratings between the coders was good (Cohen's Kappa between 0.708 and 0.995).

5.4 Analysis of the Data

Item response theory was used in order to analyze the data. This method was chosen because it made it easier to link the different booklets.[3] The method led to comparable parameters, indicating a person's ability at the different points of measurement. For coping with the requirements of different classes of items (i.e., items covering the whole modelling process and items focusing only parts of this process) a multidimensional Rasch model including subdimensions was used (cf. Brandt 2008). This model does not only estimate individual person parameters, indicating the overall modelling ability of a person, but also parameters which represent an individual's strengths and weaknesses in the implemented sub-dimension, i.e. sub-processes.

The goodness-of-fit statistics for our estimation of this Rasch model showed that one of the 36 items did not fit the model with respect to a reasonable goodness-of-fit value (mean square MNSQ $\leq$ 1.3, Wright and Linacre 1994). Thus, one underfitting item had to be excluded. In addition, five cases had to be excluded because of a failure of convergence for their response patterns (the reported sample size accounted for that loss already). For estimating the person parameters we used the WLE-method (weighted-likelihood estimates) as the most accepted method (Rost 2004, p. 314). The reliability reported for the main dimension of this multidimensional Rasch model, i.e., the dimension measuring the overall modelling competency was 0.66.

6 First Results

At this moment, we only have preliminary results of this study. They suggest that learning in the KOMMA environment was dependent of the specific forms of implementation. Since this article is mainly concerned with the structure of modelling competency and the suitability of heuristic worked-out examples for fostering this competency, we will concentrate on data from the computer-based geometry environment, which provided students with a specific program for their work and only allowed low self-regulation of the process.

6.1 Preliminary Statistical Analysis

An alpha level of 0.01 was used for all statistical analyses. As an effect-size measure, we used partial η^2 and Cohen's d (Bortz 2005, p. 145). The person parameters of the KOMMA sample were scaled to a mean of 50 and a standard deviation of 10 using a linear transformation. Table 1 shows the mean and the standard deviation of the modelling competency of this sample ($N = 316$) for all three points of measurement.

The statistical analysis reveals that learning took place in the KOMMA environment. There are significant correlations between prior modelling skills and the modelling skills in the posttest ($r = 0.45$, $p < 0.001$) as well as between prior modelling skills and the follow-up test ($r = 0.36$, $p < 0.001$).

---

[3]The IRT-analysis was based on a sample of 1657 persons who took part in all tests but were assigned to different learning environments (paper and pencil material instead of computer based instruction or material asking for a higher degree of self-regulation; see above for details).

Table 1 Mean (standard deviation) of the modelling competency in pretest, posttest, and follow-up test

	Pre-test	Post-test	Follow-up-test
Mean (standard deviation)	51.36 (9.63)	54.71 (9.93)	52.10 (9.65)

In order to test the hypothesis that learning with heuristic examples was an effective way to foster modelling competency, the person parameters of pre-test, post-test and follow-up test were contrasted by an ANOVA. We found statistically relevant variation concerning the modelling competency at the three points of measurement: $F(2.316) = 17.49$, $p < 0.001$.

The post-hoc analyses using the Bonferoni correction showed a significant development between pretest and posttest. With a value of $d = 0.46$ this effect can be categorized as a moderate one (Bortz and Döring 2002, pp. 604–605.). However, between the pretest results and the follow-up test results, there was no significant effect.

6.2 Qualitative Results of Specific Test Items

The geometry learning environment was restricted to a specific topic, namely properties of a circle, whereas the test instruments encompassed also items requiring more general geometry knowledge (see Sect. 5.3). The data presented in Sect. 6.1 were based on all test items and accordingly gave information on the general increase of competency. However, a more detailed analysis revealed differences between items asking for knowledge about a circle and its characteristics and items concerning area or circumference of other geometrical objects. In particular, most of these items show a significant increase between pretest and posttest or follow-up test. Due to restrictions in test time it was not possible to extend the multi-matrix design in such a way, that separate IRT-scalings for both content areas are possible. We will therefore restrict ourselves to a more qualitative overview on selected items.

The students were assigned to the two strands of the multi-matrix design almost equally distributed within classes. Some items were administered to all students at the same point of measurement (pretest, posttest or follow-up test). No relevant differences in solution rates could be found between the test versions on these items. We can therefore assume that the two populations assigned to different booklets did not differ in their modelling competencies. Thus, a comparison of solution rates for items that were administered to one half of the students at one point of measurement and to the other half at a later time is possible.

The test instruments encompassed 18 items with a focus on properties of a circle and its area or circumference. Five of these items were presented in pretest and posttest, three items were presented in pretest and follow-up test. Five items were presented in posttest and follow-up test but due to the multi-matrix design pretest data are not available, three items were presented only at one point of measurement. We will only consider items here that occurred in the pretest and either post- or follow-up-test.

All five items presented in pretest and posttest show an increase of correct solutions ranging from 10.3% to 25.5% (median 12.5%) percentage points. The items

As you can see in the picture, round whirlpools often have circular armrests. Give a formula that allows you to approximately calculate the length L of the armrest based on the diameter B of the pool!

Fig. 12 Item "Whirlpool" (item type 1)

The space station Mir was on its orbit for 15 years and circled the earth 86500 times in a height of 400 kilometers. The diameter of the earth is approximately 12700 km. Lee wants to estimate the distance Mir covered on one orbit. Give a drawing in which you mark the diameter of the orbit and indicate its length (use kilometer). (You do not have to calculate the length of the circuit.)

Fig. 13 Item "Mir" (item type 1)

require technical competencies as well as the construction and validation of mathematical models. Solutions rates had a higher increase when only technical knowledge had to be applied and a lower increase when competencies to build a mathematical model were needed. Thus, the item with the lowest increase, namely 10.3%, asked students to write a formula for the circumference B of a circle provided the diameter L was given. This question was embedded in the context of a circular whirlpool (Fig. 12). No calculation was needed in order to give the correct answer. In accordance to this, the rate of students who did not even try to find a solution decreased from 40.6% to 30.9%.

There was a similar trend for the item "General Sherman Tree" (see Fig. 10). This showed an increase of 12.5% correct solutions between pretest and posttest, and a similar difference in the percentage of students who did not even try to deal with it. This item asked for an explanation but did not ask for a calculation. Finally, the item with the highest increase of correct solutions was a technical one. Students were asked to calculate the area of a circle with a circumference of 12.7 centimeters. Finally, the increase of correct solutions between pretest and posttest was larger for items regarding single subcompetencies than for items covering a whole modelling process.

The three items which were presented in a pretest booklet and a follow-up test booklet give a heterogeneous picture. As described in the last paragraph, a more technical item showed an increase of 26.7% in its solution rate. This item asked students to calculate the area of a circle with a radius of 3.5 meters. The other two items showed quite similar solution rates of about 20% in both tests and similar differences of +0.5% and −3.5%. Both items involved competencies to construct (type 1) or to evaluate (type 3) a mathematical model as the example in Fig. 13 shows. Relevant and irrelevant information was provided.

7 Discussion

In summary, the findings indicate that learning with the heuristic worked examples within the learning environment KOMMA enhances students' modelling competencies. The results indicate, that heuristic worked examples are an appropriate method for supporting students initial skill acquisition in the field of modelling. As modelling activities considerably differ from proving activities, it is a major accomplishment of the present work to show the applicability and effectiveness of heuristic examples in this new field. A fundamental difference between these two fields lies in the fact that modelling tasks usually have more or less appropriate solutions, whereas proof tasks have mathematically correct solutions in a narrower sense. Modelling means that a problem solver has to employ heuristics effectively in order to arrive at the best and most accurate possible solution. Moreover, heuristic worked examples used to focus exclusively on the fostering of heuristic strategies, whereas modelling tasks required the integration of heuristic and algorithmic skills. Accordingly, we can support the theoretical implications made by Hilbert et al. (2008) from their investigations in the field of mathematical proof. They state that discovery learning or related approaches (cf., Tamir 1996) are not the only feasible ways for teaching complex skills as is usually assumed. On the contrary, example-based learning as a guided, expository learning approach—which until now was believed to be inappropriate to foster heuristic skills—is also a suitable method to attain high-level learning goals.

However, we should refrain from a too optimistic interpretation of the data. The overall test results show a significant increase of students' solution rates between pretest and posttest thus supporting the usefulness of worked examples in classroom instruction, but there is a considerable decrease of solution rates between posttest and follow-up test. Certainly, this can be regarded as a normal effect in many learning situations: Students acquire new knowledge but will forget at least a part of it over time. Nonetheless, the size of the effect in this study is relatively striking. A preliminary qualitative look on the items provides possible explanations for these results. It is evident from these data, that there are differences with respect to the specific type of items. We restricted the analysis to items, which required knowledge of the circle and its properties. For these items, we saw that modelling tasks asking explanations and interpretations were very difficult for the students before as well as after treatment. There was a large gain in the solution rates for more technically oriented modelling items. Accordingly, the learning environment might enable students to acquire more basic knowledge but will probably not support them equally in the acquisition of more demanding skills.

In particular, the increase seemed to be lowest for problems demanding the coordination of several modelling subcompetencies. Also Blum and Leiss (2007) found that technical skills were the easiest part of the modelling process for students in their DISUM study. Given the low solution rates in the pretest, the increase in the subcompetencies was nevertheless encouraging. Apart from this, the integration of the subcompetencies seemed to be a task that could not be accomplished during the short intervention. The data indicated that this integration did not occur without further instructional support after the end of the intervention.

The study gives evidence that worked examples might find their way to classroom instruction. They are suitable to support students' self-guided learning and useful

for diverse topics of the mathematics classroom. Nevertheless, there are still open questions about the sustainability of the positive results. In particular, complex skills like modelling, which encompass several quite different subcompetencies as well as content knowledge (e.g. Leiss 2007), might require a long-term intervention focusing content knowledge as well as the training of subcompetencies and the integration of these subcompetencies. All these have to be covered in a reasonable sequence (e.g., van Merriënboer and Kirschner 2007). Heuristic worked examples would certainly find a useful position in this kind of instructional sequence.

Results from classroom research (Seidel and Prenzel 2006; Hugener et al. 2008) in general and, more specifically, the DISUM project (Blum and Leiss 2006 for the comparison of two instructional styles) in the field of modelling have shown that the surface structure of instruction has only limited influence on students' competence gain. Some relevant features that should account for a supportive depth structure of modelling instruction are known. Nevertheless their implementation in interventions remains a difficult problem. Further research is needed to conceptualize reliable, theoretically underpinned teaching approaches to modelling in lower secondary school.

Regarding the structure of modelling competency, the test instruments used in the study proved to be appropriate to differentiate between theoretically derived subcompetencies. Given the complex structure of the subdimension model used in this study, an analysis of differential effects would have gone far beyond the scope of this article. Nevertheless, the first qualitative results from the intervention suggest once again that modelling competency cannot be reduced to the availability of subcompetencies. The coordination of these subcompetencies is not a trivial problem that has to be mastered by the students.

These first results reported here should be supported by further research. In particular, the specific type of knowledge and competencies to be acquired with the help of heuristic worked examples should be investigated in subsequent field studies in mathematics classrooms.

References

Aleven, V., & Koedinger, K. (2002). An effective metacognitive strategy: learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science*, *26*(2), 147–179.

Atkinson, R., Derry, S., Renkl, A., & Wortham, D. (2000). Learning from examples: instructional principles from the worked examples research. *Review of Educational Research*, *70*, 181–214.

Atkinson, R., Renkl, A., & Merrill, M. (2003). Transitioning from studying examples to solving problems: combining fading with prompting fosters learning. *Journal of Educational Psychology*, *95*, 774–783.

Blomhøj, M., & Jensen, T. (2003). Developing mathematical modeling competence: conceptual clarification and educational planning. *Teaching Mathematics and its Applications*, *22*(3), 123–139.

Blomhøj, M., & Kjeldsen, T. (2006). Teaching mathematical modeling through project work: experiences from an in-service course for upper secondary teachers. *Zentralblatt für Didaktik der Mathematik*, *38*(2), 163–177.

Blum, W. (1996). Anwendungsbezüge im Mathematikunterricht: Trends und Perspektiven. In G. Kadunz, H. Kautschitsch, G. Ossimitz, & E. Schneider (Eds.), *Trends und Perspektiven: Beiträge zum 7. Internationalen Symposium zur Didaktik der Mathematik* (pp. 15–38). Wien: Hölder-Pichler-Tempsky.

Blum, W., & Kaiser, G. (1997) Vergleichende empirische Untersuchungen zu mathematischen Anwendungsfähigkeiten von englischen und deutschen Lernenden (unpublished document). Cited according to K. Maaß (2006). What are modeling competencies? *Zentralblatt für Didaktik der Mathematik*, 38(2), 113–142.

Blum, W. & Leiss, D., Investigating quality mathematics teaching—the DISUM project. In C. Bergsten & B. Grevholm (Eds.), *Developing and researching quality in mathematics teaching and learning. Proceeding of MADIF-5 and SMDF. The fifth swedish mathematics education research seminar and swedish association for research in mathematics education* (pp. 3–16). Linköping: Sweden.

Blum, W., & Leiss, D. (2007). How do students and teachers deal with mathematical modeling problems? The example sugarloaf and the DISUM project. In C. Haines, P. Galbraith, S. Blum, & W. Khan (Eds.), *Mathematical modelling (ICTMA 12). Education, engineering and economics* (pp. 1623–1633). Chichester: Horwood Publishing.

Boero, P. (1999). Argumentation and mathematical proof: a complex, productive, unavoidable relationship in mathematic and mathematics education. *International Newsletter on the Teaching and Learning of Mathematical Proof*, pp. 7–8.

Bortz, J., & Döring, N. (2002). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Heidelberg: Springer.

Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler*. Heidelberg: Springer.

Brandt, S. (2008). Estimation of a Rasch Model including subdimensions. In M. von Davier & D. Hastedt (Eds.), *Issues and methodologies in large-scale assessments* (pp. 53–71). Hamburg: International Educational Achievement and Educational Testing Service Research Institute.

Carbonell, J. (1986). Derivational analogy: a theory of reconstructive problem solving and expertise acquisition. In R. Michalski, J. Carbonell, & T. Mitchell (Eds.), *Machine learning: an artificial intelligence approach* (pp. 371–392). San Mateo: Morgan Kaufmann.

Carroll, W. (1994). Using worked examples as an instructional support in the algebra classroom. *Journal of Educational Psychology*, *86*(3), 360–367.

Chi, M., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanation: how students study and use examples in learning to solve problems. *Cognitive Science*, *13*(2), 145–182.

Haines, C., & Crouch, R. (2006). Getting to grips with real world contexts: developing research in mathematical modeling. In M. Bosch (Ed.), *CERME 4—Conference of European research in mathematics education* (pp. 1655–1665). Guixols: Spain.

Haines, C., Crouch, R., & Fitzharris, A. (2003). Deconstruction mathematical modeling: approaches to problem solving. In Q.-X. Ye, W. Blum, K. Houston, & Q.-Y. Jiang (Eds.), *Mathematical modeling in education and culture: ICTMA 10—International community of teachers of mathematical modeling and applications* (pp. 41–53). Chichester: Horwood Publishing.

Hilbert, T., Renkl, A., Kessler, S., & Reiss, K. (2008). Learning to prove in geometry: learning from heuristic examples and how it can be supported. *Learning & Instruction*, *18*, 54–65.

Hugener, I., Pauli, C., Reusser, K., Lipowsky, F., Rakoczy, K., & Klieme, E. (2008). Teaching patterns and learning quality in Swiss and German mathematics lessons. *Learning & Instruction*, *19*, 66–78.

Jensen, T. (2007). Assessing mathematical modeling competencies. In C. Haines, P. Galbraith, W. Blum, & S. Khan (Eds.), *Mathematical modeling (ICTMA 12): Education, engineering and economics* (pp. 141–148). Chichester: Horwood.

Kaiser, G., & Sriraman, B. (2006). A global survey of international perspectives on modeling in mathematics education. *Zentralblatt für Didaktik der Mathematik*, *38*(3), 302–310.

Kultusministerkonferenz (2003). *Bildungsstandards im Fach Mathematik für den mittleren Schulabschluss*. München: Wolters Kluwer.

LeFevre, J., & Dixon, P. (1986). Do written instructions need examples? *Cognition & Instruction*, *3*, 1–30.

Legé, J. (2005). Approaching minimal conditions for the introduction of mathematical modeling. *Teaching Mathematics and its Applications*, *24*(2/3), 90–96.

Leiss, D. (2007). „*Hilf mir es selbst zu tun*": Lehrerinterventionen beim mathematischen Modellieren. Hildesheim: Franzbecker.

OECD (2003). *The PISA 2003 assessment framework: mathematics, reading, science and problem-solving knowledge and skills*. Paris: OECD (Organization for Economic Cooperation and Development).

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: recent developments. *Educational Psychologist*, *38*(1), 1–4.

Paas, F., & Van Merriënboer, J. (1994). Variability of worked examples and transfer of geometrical problem solving skills: a cognitive load approach. *Journal of Educational Psychology*, *86*, 122–133.

Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning & Instruction*, *12*(1), 61–86.

Quilici, J., & Mayer, R. (1996). Role of examples in how students learn to categorize statistics word problems. *Journal of Educational Psychology*, *88*(1), 144–161.

Recker, M., & Pirolli, P. (1995). Modeling individual differences in students' learning strategies. *The Journal of the Learning Sciences*, *4*(1), 1–38.

Reiss, K., Heinze, A., Kuntze, S., Kessler, S., Rudolph-Albert, F., & Renkl, A. (2006). Mathematikler-nen mit heuristischen Lösungsbeispielen. In M. Prenzel & L. Allolio-Näcke (Eds.), *Untersuchungen zur Bildungsqualität von Schule: Abschlussbericht des DFG-Schwerpunktprogramms* (pp. 194–208). Münster: Waxmann.

Reiss, K., Pekrun, R., Kuntze, S., Lindmeier, A., Nett, U., & Zöttl, L. (2007). KOMMA: Ein Projekt zur Entwicklung und Evaluation einer computergestützten Lernumgebung. *GDM-Mitteilungen*, *83*, 16–17.

Reiss, K., & Renkl, A. (2002). Learning to prove: the idea of heuristic examples. *Zentralblatt für Didaktik der Mathematik*, *34*(1), 29–35.

Renkl, A. (2002a). Learning from worked-out examples via self-explanations: how it can(not) be fostered. In W. Grey & C. Schunn (Eds.), *Proceedings of the 24th annual conference of the cognitive science society* (p. 49). Mahwah: Erlbaum.

Renkl, A. (2002b). Worked-out examples: instructional explanations support learning by self-explanations. *Learning & Instruction*, *12*, 529–556.

Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion* (2nd ed.). Berlin: Springer.

Schelhorn, S. E., Griego, J., & Schmid, U. (2007). Transformational and derivational strategies in analog-ical problem solving. *Cognitive Processing*, *8*(1), 45–55.

Seidel, T., & Prenzel, M. (2006). Stability of teaching patterns in physics instruction: findings from a video study. *Learning & Instruction*, *16*, 228–240.

Stark, R. (1999). *Lernen mit Lösungsbeispielen: Einfluss unvollständiger Lösungsbeispiele auf Beispiele-laboration, Lernerfolg und Motivation*. Göttingen: Hogrefe.

Sweller, J. (2003). Evolution of human cognitive architecture. *Psychology of Learning and Motivation*, *43*, 215–266.

Sweller, J., & Cooper, G. (1985). The use of worked examples as substitute for problem solving in learning algebra. *Cognition & Instruction*, *2*, 59–89.

Sweller, J., Van Merriënboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251–296.

Tamir, P. (1996). Discovery learning and teaching. In E. de Corte & F. Weinert (Eds.), *International ency-clopaedia of developmental and instructional psychology* (pp. 355–361). Oxford: Pergamon.

Treilibs, V., Burkhardt, H., & Low, B. (1980). *Formulation processes in mathematical modeling*. Notting-ham: Shell Centre for Mathematical Education.

van Gog, T., Paas, F., & Van Merriënboer, J. (2004). Process-oriented worked examples: improving transfer performance through enhanced understanding. *Instructional Science*, *32*, 83–98.

van Merriënboer, J., & Kirschner, P. (2007). *Ten steps to complex learning*. New York: Routledge.

Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen: Eine umstrittene Selbstver-ständlichkeit. In F. E. Weinert (Ed.) *Leistungsmessungen in Schulen* (pp. 17–31). Weinheim: Beltz.

Wright, B., & Linacre, M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions Contents*, *8*(3), 370. Retrieved February 11, 2009, from http://www.rasch.org/rmt/rmt83b.htm.

Zhu, X., & Simon, H. A. (1987). Learning mathematics from examples and by doing. *Cognition & In-struction*, *4*, 137–166.

Zöttl, L., & Reiss, K. (2008). Modellierungskompetenz fördern mit heuristischen Lösungsbeispielen. *Beiträge zum Mathematikunterricht*, *2008*, 189–192.