



# Towards a research base for textbooks as teacher support: the case of engaging students in active knowledge organization in the KOSIMA project

Susanne Prediger<sup>1,2</sup> · Bärbel Barzel<sup>3</sup> · Stephan Hußmann<sup>1</sup> · Timo Leuders<sup>4</sup>

Accepted: 18 February 2021  
© The Author(s) 2021

## Abstract

Engaging students in processes of mathematizing and active knowledge organization (instead of telling and rehearsing ready-made mathematics) is a key demand for high quality mathematics instruction. Although many mathematically rich context problems have been designed and published in textbooks, their impact on regular mathematics classrooms remains limited, as teachers face challenges in shaping and enacting the important phase of knowledge organization. The 15-year project KOSIMA presented in this paper was aimed at supporting teachers in really enhancing these processes by means of a textbook for Grades 5–10. The paper provides an overview of different studies conducted during the project in order to show how the project developed a research base for the support of teachers. The paper also shows that different research approaches need to be combined to establish an appropriate research base: In iterative design research cycles of topic-specific didactical design research, learning environments were designed and investigated with the aim of understanding the obstacles encountered in implementing the desired processes. In this paper, we focus on an example of one particular obstacle, namely, teachers' need of support in engaging students in processes of active knowledge organization. This obstacle was overcome by developing new types of tasks for active knowledge organization. The theoretical outcome of the design research resulted in two hypotheses for improving relevant features of the design of these tasks, which were tested in two controlled trials. Finally, a longitudinal field study of 312 students gave first indications that the classes that worked with the textbook had significantly higher learning gains than classes using other textbooks. In sum, we can show that textbooks can support changes in classroom practices when the obstacles that occur for both teachers and students are carefully investigated and treated.

## 1 Introduction

In various studies, textbooks have been shown to contribute to establishing reformed curricula in school, especially for engaging students in mathematically rich processes of inquiry (Ball & Cohen, 1996; Remillard, Harris, and Agodini 2014; Valverde et al. 2002). However, facilitating initial processes of inquiry alone is not enough, since many challenges for students and teachers have been reported, in building on the divergent knowledge construction in an initial

phase of inquiries, and *then* achieving consolidated student knowledge in an ensuing convergent phase of knowledge organization (Silver & Stein, 1996; Lobato, Clarke, and Ellis 2005; Stein et al. 2008).

We present research showing that this second phase of students' *active knowledge organization* is crucial for the effectiveness of sustainable knowledge development. However, it has turned out to be challenging for teachers to implement this phase in classrooms. In this paper, we report on a 15-year research project that aimed at supporting teachers in engaging students in these epistemic processes.

We take this practical challenge of 'engaging students in active knowledge organization' as an example, in order to argue which kind of research base is required to design textbooks that can support teachers' ambitious instructional practices. In so doing, we intend to contribute to the methodological discourse on suitable research approaches for textbook research (Heck, Chval, Weiss, and Ziebarth 2012). On the basis of our investigations, our main answer is that

✉ Susanne Prediger  
prediger@math.uni-dortmund.de

<sup>1</sup> TU Dortmund, Dortmund, Germany

<sup>2</sup> IPN Leibniz Institute for Science and Mathematics Education, Kiel/Berlin, Germany

<sup>3</sup> University Duisburg-Essen, Essen, Germany

<sup>4</sup> University of Education, Freiburg, Germany

different research approaches should be combined in order to create a suitable research base for designing textbooks that can support teachers' practices.

For this purpose, rather than presenting only one well-delineated study, we report from different studies that were conducted within a network of empirical studies. We introduce the theoretical background (Sect. 2) and the research context of the overarching long-term project KOSIMA (Sect. 3), before we present snapshots from the KOSIMA research journey with respect to the topic in view, namely, engaging students in active knowledge organization (Sect. 4). Section 5 concludes on the meta-level that, for establishing a research base for textbook design, a suitable combination of different research approaches is required.

## 2 Theoretical backgrounds

### 2.1 Textbooks and their potential for establishing ambitious teaching practices

In many countries of the world (including the authors' country, Germany), textbooks are the key source for mathematics teachers that regulate both the specification, prioritization, and sequencing of mathematical content and the kinds of tasks treated in mathematics classrooms (Keitel et al., 1980; Valverde et al. 2002). In spite of the increasing relevance of other curriculum materials, textbooks still guide many teachers' work, in particular because many other sources are more fragmented and focused on several tasks rather than the provision of coherent long-term curricula (Gravemeijer et al. 2016). Certainly, even if textbooks can only partially influence the enacted curriculum due to the active roles teachers play (Tarr, Chávez, Reys, and Reys 2006; Thompson & Senk, 2010), textbooks have been shown to “convey pedagogical messages and provide curricular environments conducive to particular teaching strategies” (Fan and Kaeley, 2000, p. 2; Rezat, Fan, and Peppin 2021).

In cases where textbook production is in the hands of commercial publishing houses that are not in contact with mathematics education research, the key role of commercial textbooks has often been problematized as hindering the quality development of instruction and retaining traditional teaching practices (Ben-Peretz, 1990).

On the other hand, textbooks playing the key role in teaching can also contribute to establishing ambitious teaching practices if the textbook design is aligned to the intended principles (Swan, 2007; Valverde et al. 2002). Indeed, evaluation studies have shown that the choice of the textbooks can have a significant effect on students' achievement (Grouws et al. 2013; van den Ham & Heinze, 2018). Therefore, we intended in our project to overcome the dichotomy between commercially produced textbooks and research-based

curriculum materials by cooperation among researchers, teachers, authors, and a commercial publisher.

With respect to the *innovative potential* of textbooks, Ball and Cohen (1996) had already problematized that most research-based curriculum materials do not sufficiently support teachers in *enacting* the intended ambitious teaching practices: “Better curriculum can only be designed if it is designed to help teachers operate more thoughtfully and effectively” (p. 8). This aspect raises the questions of how teachers deal with textbooks, and on which background they use them, either verbatim or with adaptations (Remillard, 2005).

Within the last decades, empirical studies on teachers' use of textbooks and curriculum materials have contributed to identifying typical patterns of teachers' *adaptation processes for curriculum materials* (e.g., Remillard, 2005; Sherin & Drake, 2009) and emphasized the need to value teachers' pedagogical design capacities (Brown, 2009) on their way from written to enacted curricula (Tarr et al. 2006).

Building on this richer understanding of how teachers *adapt* curriculum materials and textbooks, there is still a need to investigate how teachers can best be *supported* by textbooks, and to include results of such research in the design of textbooks (Cohen, Raudenbush, and Ball 2003; Remillard et al. 2014; Rezat et al. 2021). For this reason, we combined the insightful *descriptive research mode* of investigating teachers' practices with textbooks (for which a long research tradition and multiple methods exist; see Heck et al. 2012 and Fan, 2013) with a *constructive research mode* in action research or design research aiming to provide a research base for textbook design (for which fewer approaches have so far been developed; see Swan, 2007). This approach means rather integrating the good research tradition of describing and explaining the nature of textbooks or teachers' adaptation practices (see Fan, 2013 for an overview) into design research approaches in which intentional design and research-based improvement is also considered an important part of textbook research (which is not particularly treated in the overview of textbook research approaches by Fan, 2013).

### 2.2 Limited enactment of student-centered mathematics teaching practices and the role of the phase of knowledge organization

*Student-centered teaching practices* are a typical example of ambitious practices that have been included in curriculum materials but have not yet been sufficiently enacted in classrooms (Tarr et al. 2006; Thompson & Senk, 2010). Student-centered teaching practices are those in which students raise questions, explore situations, re-invent mathematical concepts, and discover mathematical theorems and procedures (Bruner, 1966; Freudenthal, 1973; Piaget, 1980). Maaß and

Artigue (2013) stated in their survey concerning the implementation of inquiry-based teaching, “For decades, mathematics educators have been discussing more student-centered ways of teaching, [...] They have developed theoretical constructs and materials supporting these approaches and carried out related research. *Yet, the effects on day-to-day teaching remain limited*” (p. 779).

In our project, we also adopted an inquiry-based teaching approach, more precisely we built upon traditions of Realistic Mathematics Education (RME; Freudenthal, 1973, 1991; Gravemeijer, 1999; Treffers, 1987). Freudenthal (1973) promoted the principle of guided re-invention of mathematical concepts, starting from students’ intuitive resources and developing them into more structured ideas. He already emphasized that the development of mathematical knowledge is a multi-step process of “organizing fields of experience” (p. 123). In subsequent work in the Realistic Mathematics Education Group (de Lange, 1996), examples of organizing processes are abundant, comprising “horizontal” and “vertical” mathematization (Treffers, 1987, p. 247) and their successive consolidations (more details in Sect. 3).

As Maaß and Artigue (2013) summarized, many contemporary curriculum materials and textbooks include rich open-ended tasks for inquiry. With such rich open-ended tasks, many teachers have been shown to initiate the first phase of inquiry, in which students generate multiple solutions and ideas. We regard this as a positive example of supporting teachers’ practices with adequate textbooks.

For the next phase, in contrast, authors involved in empirical classroom research have pointed out that many teachers experience challenges in maintaining the cognitive demand (Henningsen & Stein, 1997) and the richness of the mathematical discourse (Silver & Smith, 1996), and in consolidating students’ diverging ideas into “important and worthwhile mathematics” (Stein et al. 2008, p. 316), while avoiding simply ‘telling’ the ready-made mathematics (Lobato, Clarke, and Ellis 2005). Textbooks rarely contain material that supports teachers in enacting these consolidation processes while maintaining high cognitive demand. Often they only present definitions or procedures of ‘ready-made mathematics’ that do not connect to the students’ ideas or to the solutions of the inquiry phase. In many countries, teacher editions for textbooks or teachers’ manuals are available, which can give additional information on expected solutions, but often without orienting the teacher on how to use this information, or providing too-rigid scripts that lack the flexibility to adapt to students’ ideas (Cohen et al. 2003; Remillard et al. 2014).

As a consequence, Stein et al. (2008) developed professional development (PD) programs on orchestrating productive discussions for consolidating mathematical knowledge in the whole-class discussion after the inquiry phase. Their PD program addresses five teaching practices: *anticipating*

what students might invent in the lesson planning, *monitoring* students’ upcoming ideas during the inquiry phase, *selecting* those students’ strategies to be presented in the whole-class discussion, *sequencing* the presentations from the less fruitful to the most developed, and *connecting* all students’ ideas to regular mathematics. Stein et al. (2008) showed that these practices (in particular the *anticipating* and *monitoring* practices) can support the orchestration of the whole-class discussion, and teachers can learn these practices in a long-term PD program.

Since long-term PD programs are still rare in many countries (including Germany), and considering the textbooks’ central role in preparing and enacting teaching as stated above, we started our design research project with the ambition that textbook tasks and teachers’ manuals should also support teachers in the phase of knowledge organization rather than only in the inquiry phase. The design challenge for which we intended to gain a research base (Kieran et al. 2015) was therefore to develop task-based support so that more teachers might be able to enact student-centered teaching practices throughout *both phases*.

### 3 Research context and design research product

Section 3.1 briefly presents the research context of the 15-year design research project KOSIMA. Section 3.2 presents the basic design principles of the main design product, the textbook *Mathewerkstatt* and its teachers’ manual. The textbook addresses the middle school curriculum from Grades 5–10 for medium-tracked schools in Germany, called *Realschule* or *Gesamtschule* (comprehensive schools) in German. Each book was designed for one school year and consists of 8–10 teaching units.

#### 3.1 Research context of the project KOSIMA

The KOSIMA project was led by the four authors of this paper from 2005 to 2019. We collaborated with a commercial publisher, 29 textbook authors (experienced teachers and teacher educators), and 13 PhD students in design research teams. From 2012 onwards, we also qualified 34 KOSIMA PD facilitators. Establishing a research base for the textbook design activities relied mainly on the following three research approaches:

- (1) Each teaching unit for Grades 5–8 (and selected ones in Grades 9 and 10) was collectively designed and tested in *pragmatic action research* processes. Collaborating teachers tried out the drafts of the units in two to five classrooms and reflected on their experiences together with the researchers. Selected videos



**Fig. 1** Design principles of the *Mathewerkstatt* in different epistemic phases, showing the focus of this paper on active knowledge organization

and written products were considered with respect to students' and teachers' challenges, and were used to revise the teaching units. The goals of these processes were to improve the designs and gather examples from students' work for the teachers' manual as suggested by Ball and Cohen (1996).

- (2) Additionally, about one fifth of all teaching units were investigated in depth in *topic-specific design research* processes over three to five design experiment cycles in order to gain deeper insights into typical learning pathways, effects of design elements, and conditions of success for their functioning. Design research combines the following two goals: (a) improving the designs with (b) deepening the understanding of topic-specific teaching-learning processes, including empirically grounded development of local instruction theories (Gravemeijer & Cobb, 2006; Hußmann & Prediger, 2016). Within this research approach, we learned the most about major challenges for teachers, and their backgrounds, and about the potential of certain design features for overcoming them. (Topic-specific publications include Philipp, 2013, for inquiry in prime numbers; Prediger & Zwetzscher, 2013, for algebraic expressions; Prediger & Schnell, 2014, for probability; Glade & Prediger, 2017, for fractions; Büscher, 2018, for statistics; Hußmann & Prediger, 2016, for exponential functions; Schindler, 2014, for negative numbers; and Hußmann et al. 2019, for decimal numbers.)
- (3) For some of the teaching units, (*quasi-*)*randomized controlled trials* were conducted (e.g., Ganter, 2013; Loibl & Leuders, 2019; Philipp, 2013; Prediger & Wessel, 2018) with the goal of providing empirical evidence for efficacy of these teaching units and for testing hypotheses on critical design features. A large *field study* for testing overall effectiveness was conducted over two years (see Section 4.5).

### 3.2 Design principles for the textbook *Mathewerkstatt*

The design outcome of the KOSIMA project was the textbook *Mathewerkstatt* (*Math Atelier*) that was awarded Textbook of the Year in the STEM category in Germany in 2018.<sup>1</sup> The textbook comprises a student book, a teachers' manual, and many further sources such as digital applets,

practicing materials, and formative assessments tasks. As the survey by Kieran et al. (2015) revealed, design principles underlying the task design can stem from multiple sources, didactical approaches, and theoretical frameworks.

The instructional design of our textbook *Mathewerkstatt* was mainly inspired by two didactical approaches, namely, RME (Freudenthal, 1991; Gravemeijer, 1999; Treffers, 1987) and sequentially guided discovery learning with productive failure (Kapur, 2010; Loibl & Leuders, 2019).

Drawing upon RME, the design of the teaching units was based on well-established design principles (Fig. 1) that can be considered the core of RME:


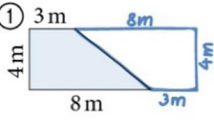
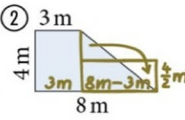
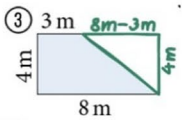
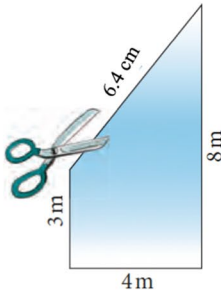
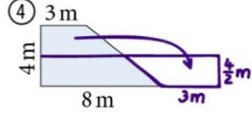
(DP1) *Develop students' conceptual understanding*: Focus on students' construction of meanings for mathematical concepts and operations and on drawing connections between knowledge elements (Hiebert & Carpenter, 1992; Wagenschein, 1977).

(DP2) *Establish and maintain high cognitive demand and active epistemic processes*: Initiate rich cognitive processes and inquiries (Barzel et al. 2013; Foster, 2018; Freudenthal, 1991; Winter, 1989; Wittmann, 1992).

(DP3) *Use multiple strategies, approaches and representations*: Provide learning opportunities in multiple representational modes and approaches and strive for students' flexible use of strategies and approaches (Duval, 2006; Durkin, Star, and Rittle-Johnson 2017; Selter, 1998; van den Heuvel-Panhuizen, 1996).

Figure 2 shows an example from Grade 8 with two tasks from a chapter on areas of geometric figures and equivalence of fractions (more details in Prediger & Zwetzscher, 2013). Rather than offering a ready-made formula for the area of the trapezoid, students are asked to find their own expressions for describing the area. The task engages students actively in inquiry processes (DP2) with the aim of developing conceptual understanding of the formula, and expressions in general (DP1). The comparison of different solutions in the second phase of active knowledge organization also provides learning opportunities for connecting multiple representational modes (DP3), which are assigned

<sup>1</sup> Textbook Award <http://www.gei.de/en/awards/textbook-of-the-year/award-winners/2018.html>.

First phase: Inquiry processes	Second phase: Active knowledge organization processes
<p>The window constructor often needs to calculate the glass area of unusual windows, as in this roof hatch.</p>	<p>The four friends found different ways of transforming the trapezoid into rectangles. <math>(3 + 8) \times 4 \div 2</math></p>
<ul style="list-style-type: none"> <li>How can he know how to calculate the area?</li> <li>Try to decompose the figure and find an expression for calculating its area.</li> </ul>	<ul style="list-style-type: none"> <li>Who has drawn which figure?</li> <li>Which expression was written by whom? Which expression is missing?</li> </ul>
	<p>① </p> <p>② </p> <p>③ </p>
	<p>④ </p> <p>I have split it twice and moved one part.</p> <p>I complemented something and had to subtract it.</p> <p>I doubled it and had to halve afterwards.</p> <p>I split it and moved one part.</p>
<p>Hint: Can you trace back this unusual figure to well-known figures?</p>	<p>Copyright Cornelsen Verlag 2015</p>

**Fig. 2** Example tasks for inquiry processes and active knowledge organization processes from *Mathewerkstatt 8* (detailed in Prediger & Zwetzscher, 2013)

to different recurring textbook characters. This design feature, called “person presentation,” has been used in many Japanese textbooks (Rittle-Johnson, 2019).

The RME approach and its existing realizations in textbooks (e.g., *Rekenen en wiskunde*, Gravemeijer & van Galen, 1984) have revealed many convincing ideas for the design of tasks for the mathematization and inquiry phase. These comprise tasks for horizontal mathematization starting from everyday context problems and leading to re-inventing mathematical concepts or discovering structures (as in the trapezoid example in Fig. 2), but also tasks for vertical mathematization that lead to abstraction and generalization (Treffers, 1987, p. 247). For example, the teaching unit from which the tasks in Fig. 2 stem continues on to the discovery of the concept of algebraic equivalence from which the transformation of expressions is later derived (Prediger & Zwetzscher, 2013).

For the third phase, practicing with intelligent exercises, mathematics education research has already developed important ideas for task design (Foster, 2018; Wittmann, 1992), so this third phase is not a focus in this paper.

In contrast, the realization of the design principles DP1-DP3 required designers’ particular attention with respect to the *second phase* of knowledge organization. This second phase is thus the focus in this paper (see Fig. 1).

Although Freudenthal (1991) and Stein et al. (2008) have already emphasized the need for knowledge organization after mathematization or inquiry, this phase has not received much attention in task design within the field of mathematics education research (Barzel et al. 2013).

In instructional psychology, the KOSIMA approach of “sequentially guided discovery learning” (Loibl & Leuders, 2019) has been investigated in the instructional framework of *problem-solving before instruction (PS-I)* (Loibl et al. 2017; see also productive failure, Kapur, 2010), to which the *Mathewerkstatt* design refers. In this strand of research, instructional designs are investigated that are structured in the following two steps: Students generate diverging or multiple solutions in an inquiry (discovery) phase. These solutions then are consolidated in a second phase of knowledge organization. The effectiveness of this instructional design, compared to a design in which instruction precedes problem solving, has been replicated in many independent studies (for a review, see Loibl et al. 2017).

So far, the phase of knowledge organization has been treated in PD mainly by enhancing teachers’ practices in preparing and conducting a whole-class discussion after the inquiry phase (Stein et al. 2008). In our project, we aimed at supporting teachers by also including textbook tasks in this phase. Beyond engaging students in more than dealing with

multiple solutions (collecting), KOSIMA aimed at engaging students in the following main epistemic processes:

- collecting, reflecting, and structuring singular and divergent ideas and strategies and connecting them to each other (*collecting* and *systematizing*);
- transforming intuitive ideas and strategies into regular and consolidated mathematics (Brousseau, 1997 calls this phase the “institutionalization”; we call it *regularizing*); and
- writing down the learned aspects in a form that is accessible for students some weeks later (*preserving*).

#### 4 Snapshots from the research journey of the project KOSIMA with respect to active knowledge organization

The research focus in the second phase emerged in early action research activities during the first years of the KOSIMA project. They are presented in a narrative style in Sect. 4.1 before we turn to more systematic research with a more rigid presentation in Sects. 4.2, 4.3, 4.4 and 4.5.

##### 4.1 Early action research results: inventing tasks and knowledge storages for supporting teachers

In our early action research processes (from 2005 to 2008), several teaching units in Grades 5 and 6 were piloted in teaching experiments. Whereas the inquiry phase mostly worked as expected, the observations of the enacted teaching practices in the phase of knowledge organization resonated with findings about teachers’ challenges documented in the literature (see Sect. 2.2). We identified four typical teaching practices:

- Only a very small number of expert teachers were able to facilitate whole-class discussions for *collecting* and *systematizing* students’ ideas and for *regularizing* them into the regular mathematical concepts. These teachers produced ‘shared knowledge storages’ in teacher-led but student-centered discussions for *preserving* the knowledge starting from students’ ideas.
- Many teachers *collected* students’ ideas and then *presented* the ready-made mathematics on the blackboard, without connecting both. By asking students to copy ready-made mathematics, these teachers succeeded in producing unified knowledge storages for the long-term *preservation* of knowledge, but at the expense of low cognitive demand. Moreover, for many students simply copying the texts did not result in understanding the pre-formulated knowledge.

- Very few teachers encouraged students to produce individual ‘knowledge storages’ in ‘inquiry diaries’ and scaffolded the processes of regularization by intensive individual, student-centered feedback. These teachers achieved *regularization* and *preserving*, but at the expense of an enormous workload while commenting on successive versions of the individual knowledge storage for each student.
- Some teachers were influenced by the idea of ‘not telling’, in other words, avoiding teacher-led instruction (Lobato et al. 2005) in such a strong way, so that all students’ ideas (whether appropriate or not) were merely *collected* and presented but left with no mathematical comments or discourse. These teachers succeeded in appraising students’ creativity, but no *systematizing* and *regularizing* took place, as the ideas remained in students’ original versions without correction.

The four teaching practices identified in the organization phase take different stances in the trade-off between (a) students’ active participation and (b) convergence towards regularized mathematics (without enormous workload for individual feedback).

In intensive action research processes in cooperation with the experimenting teachers (from 2007 to 2010), we developed means of support for teachers to cope with these trade-offs, so that less experienced teachers could also handle the first teaching practice. Therefore, support means for different epistemic processes (*collecting*, *systematizing*, *regularizing*, and *preserving*) were invented, as follows:

- (1) Support for collecting and systematizing by exploration tasks about textbook characters’ thinking

The classroom observation revealed that only those teachers who mastered Stein et al.’s (2008) five practices were able to handle the systematizing process in whole-class discussions. In order to support these practices, we addressed three of them in the task design: After the open-ended exploration task (e.g., Task 1 in Fig. 3), we included more narrow tasks that *anticipate*, *select*, and *sequence typical students’ strategies* and presented them as the strategies of four recurring textbook characters (e.g., Task 2 in ‘person presentation’ shown in Fig. 3).

Observations in later teaching experiments showed that these tasks relieved teachers’ lesson planning as the students’ *anticipated* intuitive strategies, and ideas were already provided by the textbook and explained in the teachers’ manual. For some teachers, this support was sufficient to enable them to *monitor* their own students’ upcoming strategies and continue with their students’ authentic work. The other, less flexible teachers received further support by working with the *pre-selected* and *sequenced* examples from the textbook

**Task 1: Open Competition Context Problem**

Two teams play in a competition where each player attempts to score a goal. Each player has only one attempt. Compare the scores of the two groups in the table.

	Goals for trash ball	Goals for ring toss
Girls' team (5 girls)	X X X	X X
Boys' team (10 boys)	X X X X X	X X X X

a) Who wins in trash ball: girls or boys? And in ring toss?  
 b) Discuss your different strategies for comparison. Which strategy is fair, and which one is less fair?

**Teachers' Manual: Students' individual comparing strategies for Task 1**

Non-adequate strategies focusing only one part	Partially adequate strategies with additive structuring of the situation	Adequate strategies with multiplicative structuring of the situation
<ul style="list-style-type: none"> <li>compare absolute numbers of hits</li> <li>compare absolute number of attempts</li> </ul>	<ul style="list-style-type: none"> <li>compare absolute number of failed shots, e.g., "The girls have only lost two shots, so they are better"</li> <li>transform by adding the same number for hits and attempts, e.g., "If the girls had 10 attempts, they would score 5 more, so 8"</li> </ul>	<ul style="list-style-type: none"> <li>compare failed shots relative to attempts</li> <li>compare hits relative to attempts (fractions), e.g., compare visually as in Task 2</li> <li>transform by multiplying hits and attempts by same factor, e.g., "double attempts, double hits"</li> </ul>

**Task 2: Comparing fictitious students' strategies in person presentation**

Till and Ole compare differently for the trash ball goals in Task 1.

The girls scored only three goals but the boys scored five goals.

But if you draw the bars of equal length, then it's not looking good for the boys.

a) Compare Till's and Ole's strategies.

- What is Till's argument and why?
- What is Ole's argument and why?
- Which strategy is fairer and why?

b) Apply both strategies also for the goals for ring toss for Task 1.

Copyright Cornelsen Verlag 2013

**Fig. 3** Support for anticipating, selecting, and sequencing students' strategies included in the second task, in this case visual comparison of fractions (translated and shortened from *Mathewerkstatt* 6, Prediger et al. 2013, p. 46)

characters. However, the fifth practice, *connecting* all students' ideas to the regular mathematics, was still challenging for these teachers.

(2) Support for preserving by pre-structured knowledge storages

Preserving consolidated knowledge is crucial for students' long-term memory. Simply copying the blackboard guarantees a correct entry in knowledge storage but lacks high cognitive demand. On the other hand, students' free writing requires too much individual feedback. The trade-off between activation and convergence can be found more easily in pre-structured knowledge storages that scaffold to writing an understandable (and mathematically correct!)

entry while still engaging students in reflecting on it. Figure 3 provides an example.

(3) Support for regularizing and preserving using 'active organizing tasks'

Whereas the first phase of inquiry and the third phase of practice are traditionally supported by tasks, the second phase of active knowledge organization is usually conducted without task support. Encouraged by good experiences with *systematizing tasks* in person presentation (as in Fig. 3), we also developed tasks for the *regularizing* and *preserving* processes.

The design of active *organizing tasks* focuses the acquisition activities to be initiated in students' minds. While still striving for cognitive activation, the task design should be

**Task 3 (Active organizing task)**

As discussed in Task 2, Ole and Till compared two different strategies.

a) Fill in the pre-structured knowledge store to describe the two strategies.

b) Ole compares the frequency of goals *relative* to the number of attempts. Till compares *absolute* frequencies. Explain “absolute frequency” and “relative frequency” using the sentence fragments below and explain their meaning.

If you consider only the goals,

This means, the whole number of attempts is ignored.

If you also consider the attempts,

This means, the number of goals is related to ALL attempts.

then you compare by absolute frequency.

then you compare by the share (the relative frequency).

c) Compare your solutions for b) and bring them into your knowledge storage.

**Pre-structured knowledge store**

**Strategy 1**

Girls' goals for trash ball  
x x x

Boys' goals for trash ball  
x x x x x

Here, the goals are compared **without** considering the whole number of attempts

Absolute frequency:  
If you consider only the goals, then you compare by absolute frequency. This means, the whole number of attempts is ignored.

**Strategy 2**

Girls' goals for trash ball  
[red bar] [ ] [ ] [ ] [ ] [ ]

Boys' goals for trash ball  
[blue bar] [ ] [ ] [ ] [ ] [ ]

Here, the goals are compared **with** considering the whole number of attempts.

Relative frequency:  
If you consider also the attempts, then you compare by the share (the relative frequency). This means, the number of goals is related to ALL attempts.

Copyright Cornelsen Verlag 2013

**Fig. 4** Active organizing task structuring the regularizing and preserving process (Prediger et al. 2013, p. 52)

more convergent than for inquiry tasks, in order to target consolidated knowledge for all students. Figure 4 shows the active organizing for Task 3 (which follows Tasks 1 and 2, shown in Fig. 3), which *preserves* the knowledge constructed in Task 2 in the pre-structured knowledge storage and *regularizes* the fictitious students' strategies into the mathematical concepts of absolute and relative frequency with a strong scaffold.

An adequate balance between convergence and activation depends on the concrete topic and type of knowledge (Barzel et al. 2013). For each knowledge element in students' target knowledge, the KOSIMA design teams developed repertoires of acquisition activities in this range between convergence and activation. Figure 5 shows examples of ranges for three different knowledge elements, namely, the instantiation of concepts, explicit statement of mathematical definitions or theorems, and matching of representations. The figure shows visually that increasing activation always goes along with decreasing convergence and vice versa. For example, the organizing task in Fig. 2 initiates an acquisition activity for connecting multiple representations with the highest convergence but least activation. The cognitive demand is still high due to the complexity of the content expressions for trapezoid areas.

Later action research cycles revealed that with these structured repertoires of organizing tasks, more teachers succeeded in engaging students in active knowledge organization processes, without risking too high divergence. A suitable activity structure is, for example, think-pair-share: Students first reflect alone, then share their ideas in pairs,

and then a whole-class discussion makes sure that all students fill the knowledge storage correctly.

Once the repertoire of organizing tasks was developed and categorized (see Barzel et al. 2013), the design process for the next teaching units became easier and also allowed for more rigid and systematic research activities. Extending the research base for supporting teachers in initiating active knowledge organization was still necessary for the following purposes:

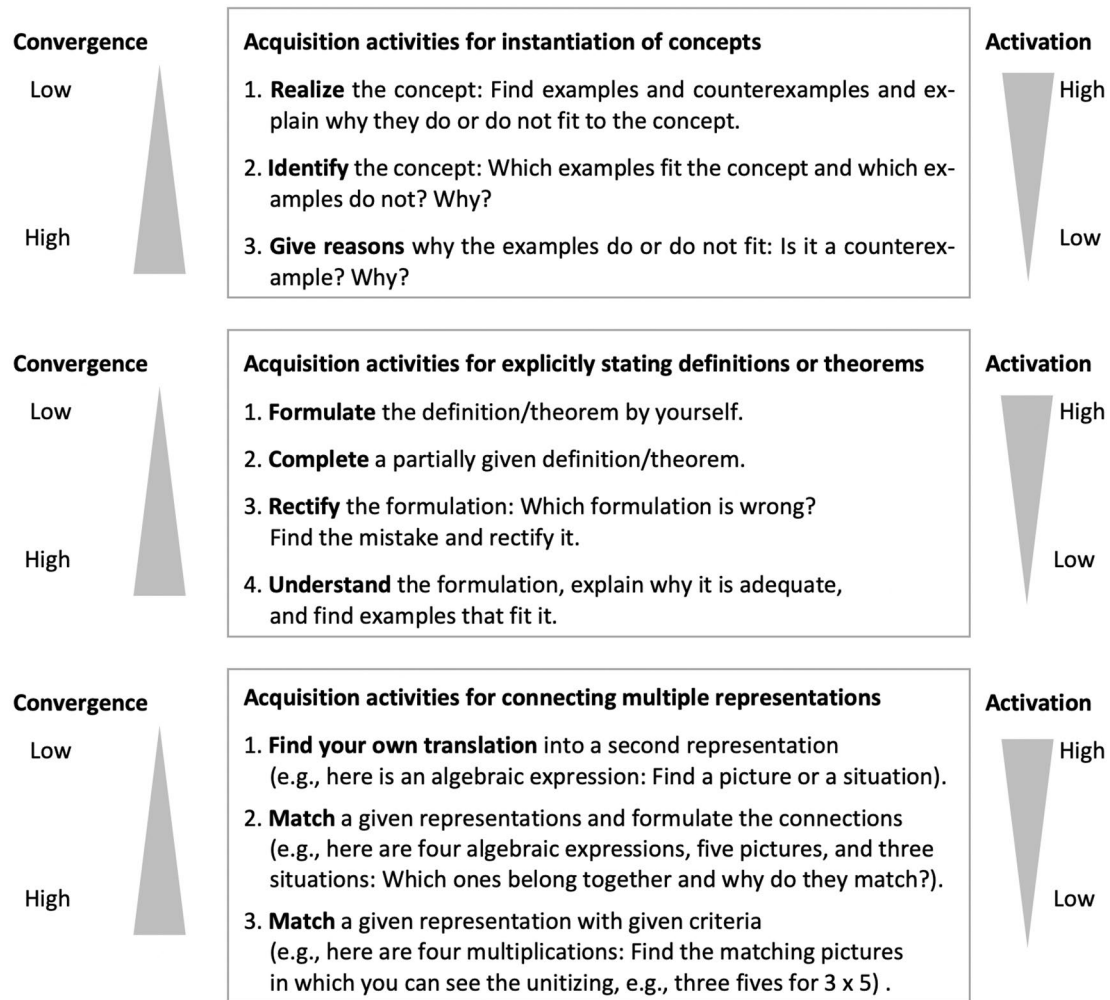
- assessing the depth of the regularizing processes for difficult mathematical topics and optimizing the topic-specific focus of organizing tasks,
- generating hypotheses for conditions of success (in an empirically grounded way), and
- corroborating these hypotheses by testing them in controlled trials.

#### 4.2 Identifying deeper challenges and generating hypotheses by design research: the need to scaffold the verbalization of meanings and structures

With the developed design features at hand (see Sect. 4.1), further design research studies were conducted for in-depth investigations of students' processes in active knowledge organization.

In a design research methodology, Glade and Prediger (2017) conducted design experiments with 18 sixth graders in pairs in order to investigate their learning pathways from a graphical part-of-part determination in area models via





**Fig. 5** Repertoire of acquisition activities in the balance between convergence and activation (Barzel et al. 2013)

progressive schematization of their graphical strategies to the discovery of the procedural rule (see Fig. 6).

The qualitative in-depth analysis of 760 min of video material revealed that progressive schematization is a vulnerable process in which students tend to jump from the graphical strategy to the procedural rule without understanding the connection between the two. Developing conceptual understanding for multiplication of fractions, however, requires more than jumping between two strategies: It requires internalizing the graphical strategies and verbalizing the underlying structures and meanings.

The analytic findings contributed to the local instruction theory about progressively schematizing multiplication of fractions (more details in Glade & Prediger, 2017). Rather than conceptualizing progressive schematization only as loss of external representations, the external representations are semiotically contracted by successive compacting of the involved structures and concepts. These processes of


compaction require explicit verbalization. Developing students' language about fractions must include developing a language for explaining meanings; in this case, phrases such as 'the part can be referred to different wholes' are more crucial than 'numerator times numerator'. Those students who verbalized the inherent multiplicative structures of the three rectangles (see Fig. 6) could also justify the procedural rule in the area model.

One immediate practical consequence for the design was that the pre-structured knowledge storage had to include scaffolds for structuring and verbalizing the arguments (see Fig. 7).

Even if the analytic procedure and this articulation of results were at first topic specific, design experiments on other topics allowed replication of these findings on the relevance of verbalizing inherent structures and meanings for successful connections of representations (e.g., algebraic

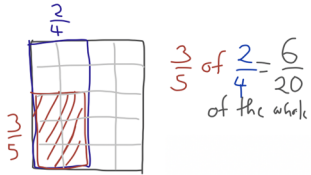
**Initial context problem: How many children go to school?**

In Chad, only one half of the children go to school and visit Grade 1. The remaining children have to work. Until Grade 5, only half of the school children stay in school.



a) Draw a rectangle to represent the situation.  
 b) What is the part of children staying in school until Grade 5?  
 c) In Pakistan,  $\frac{1}{3}$  go to school, thereof  $\frac{1}{5}$  stay until Grade 5. In Uganda,  $\frac{2}{3}$  go to school, thereof  $\frac{3}{4}$  stay.

**Horizontal mathematization into area model: graphical strategy for determining  $\frac{3}{5}$  of  $\frac{2}{4}$**



**Progressive schematization into procedural rule**

$$3 \times 2 \text{ of } 5 \times 4 = \frac{3 \times 2}{5 \times 4}$$

Copyright Cornelsen Verlag 2013

Fig. 6 From initial context problem via graphical strategy to the procedural rule (Glade & Prediger, 2017)

expressions, Prediger & Zwetzscher, 2013; and percentages, Prediger & Pöhler, 2015).

Abstracting from the concrete topic of fractions and expressions for areas, this study extended our general understanding of the epistemic process of regularizing.

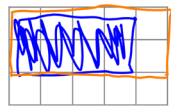
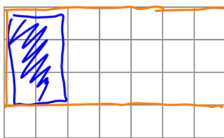
Connecting students' intuitive ideas to the regular mathematical concepts and procedures requires a *mathematical discourse about structures and meanings* that might be suitably scaffolded by given meaning-related phrases, such as the explanations given by the four textbook characters in the organizing task in Fig. 2. This result led to Hypothesis 1:

**Hypothesis 1** Regularizing from students' intuitive ideas to the regular mathematical concepts and procedures requires the permanent verbalization of meanings and structures. Teachers and students need scaffolds for this verbalization, namely, structural scaffolds and verbal scaffolds with meaning-related phrases.

### 4.3 Exploring and validating Hypothesis 1: verbalizing meanings and structures

In further design research studies, Hypothesis 1 was explored in order to gain a deeper understanding of its background (e.g., Prediger & Pöhler, 2015, dealing with percentages). In each of the topic-specific design research projects, the verbalization of meanings and structures (a) appeared as a crucial step in the process of consolidating conceptual knowledge; (b) revealed massive challenges for students, in particular for students with limited academic language proficiency; and (c) was also challenging for teachers to facilitate.

Prediger and Pöhler (2015) investigated teachers' scaffolding practices in a case study. They identified the necessary language learning trajectory that teachers' scaffolding

Knowledge storage: How to determine parts of parts		
	Example: $\frac{4}{5}$ of $\frac{2}{3}$	Own example: $\frac{2}{7}$ of $\frac{4}{5}$
I determine the part of part in the area model	 <p>2 out of 3 4 out of 5 8 shaded out of 15 in the whole rectangle</p>	 <p>3 out of 4 2 out of 7 6 out of the whole, 28</p>
The whole rectangle area consists of:	15 squares, because $5 \times 3 = 15$ .	28 squares = $4 \times 7$
The shaded part consists of:	8 squares, because $4 \times 2 = 8$ .	6 squares = $3 \times 2$
Thus, the part of part of the whole rectangle is	$\frac{8}{15}$ . $\frac{\text{part}}{\text{Whole}} = \frac{4 \times 2}{5 \times 3} = \frac{8}{15}$	$\frac{\text{part}}{\text{Whole}} = \frac{6}{28} = \frac{2}{7} \text{ of } \frac{3}{4} = \frac{2 \times 3}{7 \times 4} = \frac{6}{28}$
In total, I can calculate easily:	Thus, $\frac{4}{5}$ of $\frac{2}{3} = \frac{4 \times 2}{5 \times 3} = \frac{8}{15}$	$\frac{2}{7} \text{ of } \frac{3}{4} = \frac{2 \times 3}{7 \times 4} = \frac{6}{28}$

Copyright Cornelsen Verlag 2013

Fig. 7 Pre-structured knowledge storage: scaffolds for verbalizing the inherent structures and meanings

moves should follow in order firstly to connect students' everyday language resources to the meaning-related language for explaining meanings of concepts in the classroom discourse and then connect these phrases to the language of formal and symbolic mathematics.

Based on the identified conditions for successful and focused scaffolding, teaching units for fractions and percentages were refined with respect to this sequence of language learning opportunities. The qualitative analysis of learning processes showed that when teachers engaged students in discourses about the meaning of the part-whole relationship, students' regularizing processes became deeper (Prediger & Pöhler, 2015).

Based on these qualitative insights, Hypothesis 1 was tested in a quasi-randomized controlled trial on the topic of fractions, conducted with 343 seventh graders. Students in the language-responsive intervention acquired significantly more conceptual understanding than those in the control group receiving ordinary 'business as usual' teaching ( $F_{\text{time}} = 272.97, p < 0.001, \eta^2 = 0.45, F_{\text{time} \times \text{group}} = 22.57, p < 0.001, \eta^2 = 0.12$ ). The language-responsive intervention specifically focused on establishing a meaning-related discourse and on constantly verbalizing the connection between graphical, symbolic, and contextual representations using rich discourse practices. When comparing two versions of language-responsive support, engaging students in rich discourse practices of explaining meanings tended to be more effective than focusing on vocabulary (Prediger & Wessel, 2018). This quantitative finding is in line with qualitative findings from other researchers (Moschkovich, 2015).

Although the cited studies are still limited in scope and number of topics, they provide the first empirical evidence for Hypothesis 1.

#### 4.4 Validating Hypothesis 2 in a randomized controlled trial: active elaboration on errors

As described above, the *regularization process* requires posing and maintaining high cognitive demand with respect to connecting the individual solutions from the inquiry phase to the mathematically correct solutions and representations. Within the framework of *problem-solving before instruction* (PS-I, see Sect. 3.2, Loibl et al. 2017), the erroneous or partially correct solutions from the inquiry phase are considered as relevant preconceptions upon which the second phase should build, for instance, by comparing strategies. However, adequately to design the instructional task for these comparisons, the relevant cognitive processes must be initiated by adequate prompts (Durkin et al. 2017). Based on the students' solutions and phenomena from a pilot study, we formulated our second hypothesis:

**Hypothesis 2** Regularizing from students' intuitive ideas to the regular mathematical concepts requires the elaboration on correct and incorrect solutions and on the specific differences relevant for a conceptual change from the intuitive ideas to the mathematical ones. For that purpose, teachers should present, and students should use, prompts to explicitly focus on the relevant features of the solutions.

The hypothesis was tested in a randomized control trial with 200 fifth graders using a learning task on fraction comparison (similar to Fig. 3) with three intervention conditions (Loibl & Leuders, 2019): In an inquiry phase, students in all conditions first worked on an identical problem-solving activity (targeting the comparison of fractions) without prior instruction about the targeted concepts and procedures. In the subsequent organizing phase, students were given instruction on a correct strategy with an explanation (see Task 2 in Fig. 3). But how should students optimally process this example? In the different conditions, students were shown either (a) only correct solutions or (b) correct and erroneous examples, in order to activate them cognitively to construct deeper conceptual understanding. In a third condition, (c) they additionally received the prompt: 'Compare the solution ideas of Till and Ole. What did Till do wrong? What did Ole do better than Till?'

The ANOVA revealed that students profit most from being prompted to compare erroneous and correct solutions by their features (dark bars to the right of each set in Fig. 8), but not from merely being presented with the two solutions (striped bars in the middle of each set) or only the correct solution (light grey bars to the left). This effect is significant for those students whose articulated preconceptions were addressed in the organizing phase (middle set of bars; with  $F(2,61) = 5.69, p = 0.01, \eta^2 = 0.16, d = 0.87$ ; Loibl & Leuders, 2019). The process of students explicitly referring to the erroneous solution could be shown to mediate this learning gain.

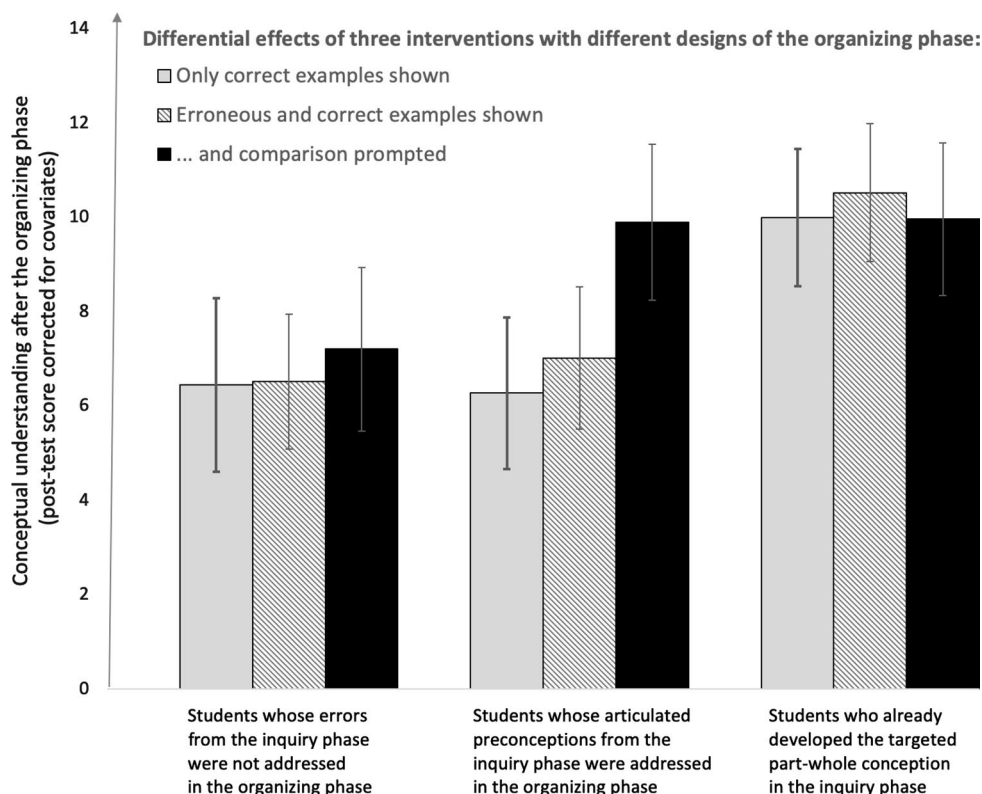
These results confirm Hypothesis 2, that the design of the organizing tasks can make significant differences.

#### 4.5 Field study testing overall effectiveness

Whereas Sects. 4.3 and 4.4 presented local empirical evidence from efficacy studies on specific hypotheses in selected teaching units, this section briefly summarizes the final field study in which the overall effectiveness of teaching with the textbook was tested during two years of schooling.

The *Mathewerkstatt* group consisted of eight classes from comprehensive schools whose mathematics teachers introduced the textbook, and these classes were accompanied by five teacher PD sessions about how to work with the textbook. One hundred and sixty-eight students from these classes completed all three tests in arithmetic at the

**Fig. 8** How should the organizing phase for students be designed? Results of a randomized controlled trial (Data from Loibl & Leuders 2019)



beginning of Grade 5, at the end of Grade 5, and at the end of Grade 6. The tasks addressed basic understanding of arithmetic topics (e.g., place value understanding, meaning of multiplication and division, and word problems with natural numbers and decimal numbers) that are fostered and used regularly throughout the two school years. The *control group* comprised eight classes from comparable comprehensive schools using their regular textbook, with 144 students from the control classes completing all tests.

As the ANOVA with repeated measures revealed, the *Mathewerkstatt* group had learning gains comparable to the control group during Grade 5 (presumably due to *Mathewerkstatt* group teachers' starting difficulties). In contrast, during Grade 6, the *Mathewerkstatt* group had significantly higher learning gains ( $F_{\text{time} \times \text{group}} = 5.34, p < 0.01$ ). However, as is expected in a field study, the effect size was small at  $\eta^2 = 0.03$  (Neumann et al. 2017). Although this result was encouraging, it was fragile as it could not be replicated in a second sample in another part of Germany.

One possible reason why the effectiveness occurred only in the second year and not in all districts might be that we measured the effects in the first year of implementation, which has often been shown to be too early, as schools

gradually introduce the new ideas (Senk & Thompson, 2003). However, a second major limitation was that no implementation control was administered, so sound proof of effectiveness has yet to be given, as the field study has provided only a first indication.

## 5 Conclusion and discussion

### 5.1 Summary and limitations of the presented snapshots from the KOSIMA project on active knowledge organization

Ambitious instructional practices, such as meaningful and inquiry-based learning, are hard to establish in mathematics classrooms (Maaß & Artigue, 2013; Silver & Stein, 1996). In this paper, we argue that careful and research-based textbook design can substantially support teachers in realizing them. Whereas various textbooks provide convincing teacher support for engaging students in the first phase of inquiry, the second phase of active knowledge organization is a good example of a teacher challenge that requires more support (Silver & Stein, 1996) but has so far been mainly tackled in professional development (e.g., Stein et al. 2008) rather than in textbooks.

Although this paper cannot account for the whole complexity of a 15-year research project, it offers snapshots into selected typical steps. These steps document how the challenge was treated by understanding it more deeply in action research processes and design research processes (Sects. 4.1 and 4.2), by designing tasks and scaffolds for students and teachers, and by conducting studies for exploring and validating hypotheses about specific design features (Sects. 4.3 and 4.4). Of course, the snapshots are all limited, supported by necessarily brief presentations of research that is documented more thoroughly in the cited papers. However, by connecting the steps to a presentation of the research journey through five studies, we intended to present a larger picture and a more complex contribution to the questions concerning how to provide teacher support for initiating knowledge organization using textbooks.

The initial *action research experiments* confirmed the existing findings about teachers' abilities to realize inquiry phases in mathematics classrooms when they work with rich open-ended context problems (Freudenthal, 1973, 1991). However, they also confirmed the observations about teachers' challenges in realizing the second phase of active knowledge organization (Lobato et al. 2005; Silver & Stein, 1996). As a consequence, the design teams constructed tasks for comparing strategies and tasks for active knowledge organization that support teachers in realizing sophisticated facilitation practices (Stein et al. 2008), more specifically for engaging students in four crucial epistemic processes: *collecting*, *systematizing*, *regularizing*, and *preserving* (Barzel et al. 2013).

The subsequent *design research endeavors* (e.g., on students' progressive schematization processes, Treffers, 1987) provided deeper insights on typical conditions of success for regularizing (Glade & Prediger, 2017; Prediger & Pöhler, 2015). Two hypotheses were formulated on how to regularize from students' ideas to regular mathematics: Hypothesis 1 focused the need for permanent verbalization of meanings and structures and Hypothesis 2 on the elaboration on correct and incorrect solutions relevant for conceptual change.

Both hypotheses were further explored qualitatively and then validated in controlled trials (Loibl & Leuders, 2019; Prediger & Pöhler, 2015; Prediger & Wessel, 2018). In both trials, the intervention groups, working with the design features hypothesized as relevant, outperformed the control group in the post-test (Sects. 4.3 and 4.4). Although these results provide first local empirical evidence, their scope must not be overinterpreted, as the findings are still tied to the very specific research contexts, the specific topics, and settings in view. Future studies should broaden their scope by transferring the studies from fractions to further topics and by further exploring the exact conditions under which they work. In particular, more thorough process studies need to be conducted in order to capture teachers' enactment.

In a huge textbook design research project that encompasses a whole middle school curriculum from Grades 5 to 10, a highly controlled research project can cover only a small part of the material. In our case, only about 20% of the teaching units were investigated in design research studies, and less than 5% in controlled trials. Future research should continue to investigate other parts of the curriculum, but completeness can never be reached.

For Grades 5 and 6, at least, a field test was able to provide some first empirical indications for the overall effectiveness of the textbook. Although this result was encouraging, it was still fragile and not yet replicable in a second sample. Due to the methodological limitations of big-sample drop-out rate and missing implementation control in teachers' enactment, these findings must be treated with caution, and further studies are necessary before claiming the textbook's effectiveness. A further limitation is that in a big field test for a two-year curriculum, which design features were particularly crucial is no longer clear. Future implementation research will be necessary with systematic variations in the design in order to disentangle the effects of different design features.

Additionally, the field test revealed large differences in average learning gains of different classes. We interpret these class effects as a potential indicator for the important role of teachers' design capacities (Brown, 2009; Cohen et al. 2003), which goes along with different enacted teaching practices. Future studies will need to control for the implementation quality of the enacted curriculum (Heck et al. 2012; Tarr et al. 2006), because with the current data, we cannot test this interpretation statistically nor generate explanations for the different impact in different school systems. This will also allow the disentangling of what exactly a textbook or other curriculum material can achieve from what was determined by the professional development itself.

Although the textbook could be shown to *support* teachers, its enactment can evidently not be a substitute for teachers' expertise (Ball & Cohen, 1996; Swan, 2007). From this we have concluded that there is also a need for professional development courses. Future studies will have to investigate the interplay of textbook support and professional development in more depth (as requested by Senk & Thompson, 2003; Tarr et al. 2006).

## 5.2 Discussion on the meta-level: How to establish a research base for textbook design using different research approaches

The research journey sketched in this paper also provides an answer to the meta-level question of how to establish an *interventionist rather than descriptive research base for textbook design* (a perspective that has not been in the main

focus of existing textbook research; see Fan, 2013). No single research approach can provide a comprehensive research base because each approach has strengths and limitations. As Burkhardt and Schoenfeld (2003) have emphasized for other fields of research, the combination of different research approaches can enhance the understanding of teacher support by textbook design in the following ways:

*Classroom observation studies* can help to identify productive and less productive teaching practices and the ways teachers enact the textbook curriculum. These studies have been increasingly refined by multiple methods (Heck et al. 2012; Tarr et al. 2006). However, descriptive approaches alone cannot help to overcome teachers' challenges.

*Action research studies* can help collaboratively to develop solutions for teachers' challenges by designing elements for textbooks that can support teachers in enhancing their teaching practices. However, this approach is too limited to deepen the understanding of the process and to contribute to theory generation.

*Design research studies* combine iterative cycles of design and design experiments with deeper qualitative investigations of the generated teaching learning processes. They can provide substantial empirical insights into patterns involved in teaching–learning processes, and generate hypotheses about typical effects and conditions of success for design elements. However, design research with its qualitative methods can only establish and explore hypotheses but not validate them.

*Controlled trials* form a well-established research approach for validating hypotheses about the efficacy of design elements in textbooks. They build upon the qualitative investigation of the teaching learning processes where hypotheses were generated. As efficacy is often shown in laboratory settings with highly controlled conditions, these laboratory trials cannot yet provide ecological validity for the functioning of the design elements under realistic field conditions. For studying design elements for supporting teachers, this is a serious constraint.

*Field tests* form the research approach of choice, providing increased ecological validity by testing the textbook under field conditions with more teachers. While it is well known that field tests can achieve only smaller effect sizes, and raise many questions about comparability, they are important because they provide empirical evidence for the effectiveness of the designs.

The affordances and limitations of each of these research approaches show why their combination is so crucial for the long-term research program. This is well known in other fields of research and should (in the long run) also be established for a textbook research program

that aims at generating theory and improving classroom practices at the same time.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Ball, D. L., & Cohen, D. K. (1996). Reform by the book: What is—or might be—the role of curriculum materials in teacher learning and instructional reform? *Educational Researcher*, 25(9), 6–8.
- Barzel, B., Leuders, T., Prediger, S., & Hußmann, S. (2013). Designing tasks for engaging students in active knowledge organization. In A. Watson, M. Ohtani, J. Ainley, F. J. Bolite, M. Doorman, C. Kieran, A. Leung, C. Margolinas, P. Sullivan, D. Thompson, & Y. Yang (Eds.), *ICMI Study 22 on Task Design: proceedings of the study conference* (pp. 285–294). ICMI.
- Ben-Peretz, M. (1990). *The teacher-curriculum encounter: Freeing teachers from the tyranny of texts*. SUNY Press.
- Brousseau, G. (1997). *The theory of didactical situations in mathematics*. Kluwer.
- Brown, M. W. (2009). The teacher-tool relationship. Theorizing the design and use of curriculum materials. In J. T. Remillard, B. A. Herbel-Eisenmann, & G. M. Lloyd (Eds.), *Mathematics teachers at work: Connecting curriculum materials and classroom instruction* (pp. 17–36). Routledge.
- Bruner, J. (1966). *Toward a theory of instruction*. Harvard University Press.
- Burkhardt, H., & Schoenfeld, A. (2003). Improving educational research: Toward a more useful, more influential, and better-funded enterprise. *Educational Researcher*, 32(9), 3–14.
- Büscher, C. (2018). *Mathematical literacy on statistical measures. A design research study*. Springer Spektrum.
- Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, 25(2), 119–142.
- De Lange, J. (1996). Using and applying mathematics in education. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 49–97). Kluwer.
- Durkin, K., Star, J. R., & Rittle-Johnson, B. (2017). Using comparison of multiple strategies in the mathematics classroom: Lessons learned and next steps. *ZDM – Mathematics Education*, 49(4), 585–597.
- Duval, R. (2006). A cognitive analysis of problems of comprehension in a learning of mathematics. *Educational Studies in Mathematics*, 61(1–2), 103–131.
- Fan, L. (2013). Textbook research as scientific research: Towards a common ground on issues and methods of research on mathematics textbooks. *ZDM – Mathematics Education*, 45(5), 765–777.

- Fan, L., & Kaeley, G. (2000). The influence of textbooks on teaching strategies: An empirical study. *Mid-Western Educational Researcher*, 13(4), 2–9.
- Foster, C. (2018). Developing mathematical fluency: Comparing exercises and rich tasks. *Educational Studies in Mathematics*, 97(2), 121–141. <https://doi.org/10.1007/s10649-017-9788-x>
- Freudenthal, H. (1973). *Mathematics as an educational task*. Reidel.
- Freudenthal, H. (1991). *Revisiting mathematics education*. Kluwer.
- Ganter, S. (2013). *Experimentieren: ein Weg zum Funktionalen Denken: Empirische Untersuchung zur Wirkung von Schülerexperimenten*. Kovac.
- Glade, M., & Prediger, S. (2017). Students' individual schematization pathways: Empirical reconstructions for the case of part-of-part determination for fractions. *Educational Studies in Mathematics*, 94(2), 185–203. <https://doi.org/10.1007/s10649-016-9716-5>
- Gravemeijer, K. (1999). How emergent models may foster the constitution of formal mathematics. *Mathematical Thinking and Learning*, 1(2), 155–177.
- Gravemeijer, K., & Cobb, P. (2006). Design research from a learning design perspective. In J. Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research: The design, development and evaluation of programs, processes and products* (pp. 17–51). Routledge.
- Gravemeijer, K., & van Galen, F. (1984). *Rekenen en wiskunde [Textbook]*. Bekadidact.
- Gravemeijer, K., Bruin-Muurling, G., Kraemer, J.-M., & van Stiphout, I. (2016). Shortcomings of mathematics education reform in The Netherlands: A paradigm case? *Mathematical Thinking and Learning*, 18(1), 25–44. <https://doi.org/10.1080/10986065.2016.1107821>
- Grouws, D. A., Tarr, J. E., Chávez, Ó., Sears, R., Soriaand, V. M., & Taylan, R. D. (2013). Curriculum and implementation effects on high school students' mathematics learning from curricula representing subject-specific and integrated content organizations. *Journal for Research in Mathematics Education*, 44(2), 416–463. <https://doi.org/10.5951/jresmetheduc.44.2.0416>
- Heck, D. J., Chval, K. B., Weiss, I. R., & Ziebarth, S. W. (Eds.). (2012). *Approaches to studying the enacted mathematics curriculum*. Information Age.
- Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, 28(5), 524–549.
- Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 65–97). Macmillan.
- Hußmann, S., & Prediger, S. (2016). Specifying and structuring mathematical topics: A four-level approach for combining formal, semantic, concrete, and empirical levels exemplified for exponential growth. *Journal für Mathematik-Didaktik*, 37(Suppl. 1), 33–67. <https://doi.org/10.1007/s13138-016-0102-8>
- Hußmann, S., Schacht, F., & Schindler, M. (2019). Tracing conceptual development in mathematics: Epistemology of webs of reasons. *Mathematics Education Research Journal*, 31(2), 133–149.
- Kapur, M. (2010). Productive failure in mathematical problem solving. *Instructional Science*, 38(6), 523–550.
- Keitel, C., Otte, M., & Seeger, F. (1980). *Text Wissen Tätigkeit*. Scriptor.
- Kieran, C., Doorman, M., & Ohtani, M. (2015). Frameworks and principles for task design. In A. Watson & M. Ohtani (Eds.), *Task design in mathematics education: ICMI Study 22* (pp. 19–81). Springer.
- Lobato, J., Clarke, D., & Ellis, A. B. (2005). Initiating and eliciting in teaching: A reformulation of telling. *Journal for Research in Mathematics Education*, 36(2), 101–136. <https://www.jstor.org/stable/30034827>
- Loibl, K., & Leuders, T. (2019). How to make failure productive: Fostering learning from errors through elaboration prompts. *Learning and Instruction*, 62, 1–10. <https://doi.org/10.1016/j.learninstruc.2019.03.002>
- Loibl, K., Roll, I., & Rummel, N. (2017). Towards a theory of when and how problem solving followed by instruction supports learning. *Educational Psychology Review*, 29(4), 693–715.
- Maaß, K., & Artigue, M. (2013). Implementation of inquiry-based learning in day-to-day teaching: A synthesis. *ZDM – Mathematics Education*, 45(6), 779–795.
- Moschkovich, J. (2015). Academic literacy in mathematics for English Learners. *The Journal of Mathematical Behavior*, 40(A), 43–62.
- Neumann, R., Leuders, T., Prediger, S., Hußmann, S., Barzel, B., & Altieri, M. (2017). Wirksamkeit des KOSIMA-Unterrichtskonzept im Feld: Längsschnittliche Evaluation in Klasse 5 und 6. In U. Kortenkamp & A. Kuzle (Eds.), *Beiträge zum Mathematikunterricht* (pp. 701–704). WTM.
- Philipp, K. (2013). *Experimentelles Denken. Theoretische und empirische Konkretisierung einer mathematischen Kompetenz*. Springer Spektrum.
- Piaget, J. (1980). The psychogenesis of knowledge and its epistemological significance. In M. Piattelli-Palmarini (Ed.), *Language and learning* (pp. 23–54). Harvard University Press.
- Prediger, S., & Pöhler, B. (2015). The interplay of micro- and macro-scaffolding: an empirical reconstruction for the case of an intervention on percentages. *ZDM – Mathematics Education*, 47(7), 1179–1194. <https://doi.org/10.12973/eurasia.2015.1497a>
- Prediger, S., & Schnell, S. (2014). Investigating the dynamics of stochastic learning processes. In E. J. Chernoff & B. Sriraman (Eds.), *Probabilistic thinking* (pp. 533–558). Springer. [https://doi.org/10.1007/978-94-007-7155-0\\_29](https://doi.org/10.1007/978-94-007-7155-0_29)
- Prediger, S., & Wessel, L. (2018). Brauchen mehrsprachige Jugendliche eine andere fach- und sprachintegrierte Förderung als einsprachige? Differentielle Analysen zur Wirksamkeit zweier Interventionen in Mathematik. *Zeitschrift für Erziehungswissenschaft*, 21(2), 361–382. <https://doi.org/10.1007/s11618-017-0785-8>
- Prediger, S., & Zwetzschler, L. (2013). Topic-specific design research with a focus on learning processes: The case of understanding algebraic equivalence in grade 8. In T. Plomp & N. Nieveen (Eds.), *Educational design research: Illustrative cases* (pp. 407–424). SLO, Netherlands Institute for Curriculum Development.
- Prediger, S., Barzel, B., Hußmann, S., & Leuders, T. (2013). *Mathewerkstatt 6 [Textbook for German middle schools, Grade 6]*. Cornelsen.
- Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research*, 75(2), 211–246.
- Remillard, J. T., Harris, B., & Agodini, R. (2014). The influence of curriculum material design on opportunities for student learning. *ZDM – Mathematics Education*, 46(5), 735–749. <https://doi.org/10.1007/s11858-014-0585-z>
- Rezat, S., Fan, L., & Peppin, B. (2021, in press). Mathematics textbooks and curriculum resources as instruments for change. *ZDM – Mathematics Education*, 53(6).
- Rittle-Johnson, B. (2019). *Learning new strategies: Insights from Japanese math textbooks*. Society for Research in Child Development.
- Schindler, M. (2014). *Auf dem Weg zum Begriff der negativen Zahl. Empirische Studie zur Ordnungsrelation für ganze Zahlen aus inferentieller Perspektive*. Springer Spektrum.
- Selter, C. (1998). Building on children's mathematics. A teaching experiment in Grade 3. *Educational Studies in Mathematics*, 36(1), 1–27.
- Senk, S., & Thompson, D. (Eds.). (2003). *Standards-based school mathematics curricula: What are they? What do students learn?* Lawrence Erlbaum.

- Sherin, M. G., & Drake, C. (2009). Curriculum strategy framework: Investigating patterns in teachers' use of a reform-based elementary mathematics curriculum. *Journal of Curriculum Studies*, 41(4), 467–500.
- Silver, E. A., & Smith, M. S. (1996). Building discourse communities in mathematics classrooms: A worthwhile but challenging journey. In P. C. Elliot & M. J. Kenny (Eds.), *Communication in mathematics: K-12 and beyond* (pp. 20–28). NCTM.
- Silver, E. A., & Stein, M. K. (1996). The QUASAR project: The “revolution of the possible” in mathematics instructional reform in urban middle schools. *Urban Education*, 30(4), 476–521.
- Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning*, 10(4), 313–340. <https://doi.org/10.1080/10986060802229675>
- Swan, M. (2007). The impact of task-based professional development on teachers' practices and beliefs: A design research study. *Journal of Mathematics Teacher Education*, 10(4–6), 217–237.
- Tarr, J. E., Chávez, Ó., Reys, R. E., & Reys, B. J. (2006). From the written to the enacted curricula: The intermediary role of middle school mathematics teachers in shaping students' opportunity to learn. *School Science and Mathematics*, 106(4), 191–201.
- Thompson, D. R., & Senk, S. L. (2010). Myths about curriculum implementation. In B. J. Reys, R. E. Reys, & R. Rubenstein (Eds.), *Mathematics curriculum: Issues, trends, and future directions: Seventy-second yearbook* (pp. 247–263). National Council of Teachers of Mathematics.
- Treffers, A. (1987). *Three dimensions: A model of goal and theory description in mathematics instruction: The Wiskobas project*. Kluwer.
- Valverde, G. A., Bianchi, L. J., Wolfe, R. G., Schmidt, W. H., & Houang, R. T. (2002). *According to the book: Using TIMSS to investigate the translation of policy into practice through the world of textbooks*. Kluwer.
- Van den Heuvel-Panhuizen, M. (1996). *Assessment and realistic mathematics education*. Utrecht University Press.
- Van den Ham, A.-K., & Heinze, A. (2018). Does the textbook matter? Longitudinal effects of textbook choice on primary school students' achievement in mathematics. *Studies in Educational Evaluation*, 59, 133–140.
- Wagenschein, M. (1977). *Verstehen lehren. Genetisch, Sokratisch, Exemplarisch*. Beltz.
- Winter, H. (1989). *Entdeckendes Lernen im Mathematikunterricht*. Vieweg.
- Wittmann, E. C. (1992). Wider die Flut der bunten Hunde und der grauen Päckchen: Die Konzeption des aktiv-entdeckenden Lernens und produktiven Übens. In G. Müller & E. C. Wittmann (Eds.), *Handbuch produktiver Rechenübungen* (pp. 152–166). Klett.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.