Educational Studies in Mathematics
https://doi.org/10.1007/s10649-018-9835-2

Combining material- and community-based implementation strategies for scaling up: the case of supporting low-achieving middle school students

Susanne Prediger[1] · Claudia Fischer[2] · Christoph Selter[1] · Christian Schöber[2]

Abstract
In order to better facilitate scaling up of classroom innovations, two complementary strategies have often been discussed. The *community-based strategy* emphasizes the necessity for professional learning communities and their embedding in institutional settings. The *material-based strategy* starts from well-designed teaching materials, which are considered catalysts for bringing teaching approaches to many classrooms. The implementation project reported on in this study systematically combines both strategies and takes a third strategy into account: the systemic strategy of addressing higher levels of the school system, such as the school and district levels. The goal of the project is to help teachers to better support low-achieving students at the beginning of German secondary schools (grades 5 and 6). The results of the accompanying research in a quasi-experimental study, reported in this article, show that a combination of strategies can be effective: the participating low-achieving students had higher learning gains than a control group did. The deeper analysis provides insights into the complexities of the interplay of community aspects, institutional backgrounds, and the power of substantial teaching materials.

Keywords Implementation · Material for scaling up · Professional learning communities · Low-achieving students

It has often been emphasized that educational and didactical innovations—for example alternative teaching approaches, new curricula, and research results—do not spread

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

Claudia Fischer
cfischer@ipn.uni-kiel.de

1 Institute for Development and Research in Mathematics Education (IEEM), TU Dortmund University, 44221 Dortmund, Germany

2 Leibniz-Institute for Science and Mathematics Education (IPN), 24118 Kiel, Germany

Published online: 20 August 2018

automatically in a school context, but have to be accompanied and supported, especially when implementing at scale (Coburn, 2003; Rogers, 2003). Some teachers might not consider these innovations beneficial; others struggle to reconcile the changes with their experiences and subjective theories. The need for support is underscored by many empirical studies (Euler & Sloane, 1998; Schellenbach–Zell, Rürup, Fussangel, & Gräsel, 2008), showing that dissemination of innovations is often slow or hesitant especially when scaling up (Gräsel, Fußangel, & Pröbstel, 2006; Rogers, 2003).

Thus, scaling up is the huge challenge for all educational reforms, even if there is evidence for the efficacy of a teaching approach in laboratory and other smaller-scale studies (Adler & Jaworski, 2009; Burkhardt & Schoenfeld, 2003; Roesken-Winter, Hoyles, & Bloemeke, 2015). In order to promote the scaling-up processes, two main implementation strategies are often discussed: the strategy based on professional learning *communities* on the one hand and the strategy based on well-designed *material* on the other. This article shows how community-based and material-based strategies can be treated as complementary. The presented scaling-up project for supporting low-achieving middle school students combines both strategies, and the ongoing formative evaluation provides empirical insights into its functioning. In sum, the article has the goal of giving an existence proof for the ability to realize the combining of both strategies by presenting a possible project architecture and providing both empirical evidence for its effectiveness on students' learning gains and some deeper insights into the impact they may have on other levels.

Section 1 outlines the state of discourse on different implementation strategies. Section 2 describes the content of the project: supporting low achievers. Section 3 explains and justifies the complex architecture of the implementation project and elaborates the research context. Section 4 presents the methods of data gathering and analysis. Section 5 presents insights into the scaling-up process and its effectiveness, opportunities, and challenges. Finally, Section 6 presents a summarizing discussion and some concluding remarks.

1 Implementation strategies for scaling up

1.1 Material-based strategy: designing robust teaching units and manuals

Most teachers use resources as guidelines in practice (McDuffie & Mather, 2006), such as textbooks or teacher's manuals. Maaß and Artigue (2013, p. 784) stated that there is a growing body of research showing that teaching resources affect instructional practice and student learning. In this article, approaches that start from the assumption that well-designed teaching materials are an important catalyst for bringing teaching approaches to many classrooms are subsumed under the notion of *material-based strategy.*

Material-based implementation projects have shown that the careful design of materials can substantially support the quality and coherence of implemented teaching approaches (Burkhardt, 2006; Burkhardt & Schoenfeld, 2003; Swan, 2007). As argued by Gravemeijer, Bruin-Muurling, Kraemer, and van Stiphout (2016), this material needs to encompass complete teaching units or curricula, not only single tasks. The quality of the support increases with the research-based quality of the materials (Roesken-Winter et al., 2015): They should be empirically founded on a thorough analysis of the problems to be tackled and come with quantitative empirical evidence of their efficacy.

However, a purely material-based strategy is too limited since teachers need an environment in which they can appropriate the materials to their own contexts (Remillard, 2005). As

Maaß and Artigue (2013) summed up: "resources can be considered as an important component of dissemination, but which, when used without other strategies for dissemination, are only of limited impact" (p. 784). Thus, dissemination of materials must be accompanied by the professional development of teachers. Rather than top-down teacher training models, the material-based dissemination strategy is combined here with a community-based strategy.

1.2 Community-based strategy: establishing professional learning communities

In their research review on conditions of successful professional development of teachers, Darling-Hammond and Richardson (2009) identified six crucial aspects. It should:

- Enhance teachers' subject knowledge and the way it is shared with students;
- Help teachers understand how students learn certain concepts;
- Create opportunities for active learning through participation;
- Enable teachers to expand their knowledge, put it into practice, and jointly reflect;
- Link the curriculum, assessment, and standards to professional learning; and
- Be cooperative, collegial, intensive, and organized over a longer period of time.

These findings about conditions for successful instruction-oriented teacher training call for more collaborative and reflective settings (Krainer, 2008; Selter & Bonsen, 2018). *Community-based strategies* emphasize the necessity for stable communities of teachers jointly engaging in long-term change processes. Examples are lesson studies (Kullmann, 2012; Takahashi & Yoshida, 2004) and professional learning communities (Bonsen & Rolff, 2006; Lomos, Hofman, & Bosker, 2011). Professional learning communities are groups of teachers (mostly at the same school) who jointly deepen their knowledge and understanding of what is taught, design lessons, take actions in their classrooms, assess student results, and reflect on their actions and on student results (Selter & Bonsen, 2018). As innovative teaching practices cannot be adopted by teachers simply reading research papers or manuals, professional learning communities are important institutions where teachers can find their own ways to realize the general principles of innovative teaching approaches (Krainer, 2008).

On the other hand, purely bottom-up approaches are challenging for many teachers when there is not enough didactical orientation. That is why we follow Lachance and Confrey (2003) in their plea for combining both the material- and community-based strategies: Establishing professional communities is crucial to giving teachers the social context for systematic and continuous reflection and offers them research-based teaching materials and gives them orientation on how to start.

1.3 Systemic strategy: addressing multi-level complexities

As Llinares and Krainer (2006) pointed out, research on teachers' processes of professional change must take into account not only the *content* (here supported by the material-based strategy) and *communities* (here supported by the community-based strategy), but also the institutional *contexts* in which the teachers work. Considering a professional development program in context means taking into account the teachers' schools, school districts, and the general conditions (resources, structures, commitment, etc.). These institutional contexts span from the teacher community level and the school context to the school district level with their local authorities, as all these contexts can support or hinder the change processes in significant ways (Cobb & Jackson, 2012; Llinares & Krainer, 2006). In particular, school management has been shown to be crucial for change processes (Bryk, Sebring, Allensworth, Luppescu, & Easton, 2010; Darling-Hammond, 1997; Prenzel, Friedrich, & Stadler, 2008). Thus, it is necessary to address the communities within their specific institutional settings, which requires a multi-level structure in the project design (see Fig. 3).

1.4 Design challenge and research questions

Based on the outlined state of research, this article focuses on a general design challenge:

(DC) How can a research-based implementation and professional development program be designed that combines the three implementation strategies: material based, community based, and systemic?

Section 3 of this article will present the project architecture as an answer to this high practical-interest design challenge. Sections 4 to 6 will show the theoretical relevance of this work will be shown in the presented research project, which pursued the following three research questions:

(RQ1) What are the effects of the research-based implementation and professional development program on teachers' perceptions of materials and of their cooperation in professional learning communities?

(RQ2) What are the effects of the research-based implementation and professional development program on students' learning gains compared to a control group of students supported by teachers outside the program?

(RQ3) What are possible challenges to and advantageous conditions for the success of the program on the different levels?

2 Research-based background of the innovation in view: fostering basic conceptual understanding for low-achieving students

The research questions were posed and the general design challenge was examined in the *Mastering Math* ("Mathe sicher können") project, whose content and research-based background is described in this section. Mastering Math is a design research and professional development project of the Institute for Development and Research in Mathematics Education at TU Dortmund University (Prediger, Selter, Hußmann, Nührenbörger, and others). It is funded by the German Telekom Foundation from 2011 to 2020. Since 2015, it has been externally evaluated by IPN Kiel (Köller, Fischer, Döring, Trepke, and Schöber). The goal of the project is to develop teachers' expertise in supporting low-achieving students in German middle schools (grades 5 and 6, ages 10–12).

2.1 Research background: necessity for the intervention program and its efficacy on fostering basic conceptual understanding of low-achieving middle school students

The practical starting point of the long-term research and development project was the statistical findings of large-scale assessments showing that about 20% of German middle school students do not reach the curricular goals of middle schools (Prenzel, Sälzer, Klieme, & Köller, 2013; an internationally similar example is Maccini, Mulcahy, & Wilson, 2007). Empirical research on low-achieving students' knowledge structures showed that most difficulties can be traced back to difficulties in primary education content (Andersson, 2010; Selter, Wessel, Walther, & Wendt, 2012): In the German context, this has consisted mainly of deficits in basic conceptual understanding of the place value system and the basic operations (i.e., multiplication and division, relating representations, and word problems; see Moser Opitz, 2007).

Based on these empirical results, the pilot phase (2009–2012) focused on developing and evaluating an intervention program in 14 lessons for students to foster their basic conceptual needs. A quasi-randomized control trial involving grades 5 and 6 students ($n = 123$) provided empirical evidence on its efficacy (Moser Opitz et al., 2017). In addition, a standardized curricular test was developed and optimized psychometrically in order to select those students with basic conceptual needs in an objective, reliable, and valid way (Moser Opitz, Freesemann, Grob, & Prediger, 2016).

Based on this foundation, the Mastering Math project was initiated, with the goal of bringing the pilot intervention program from the laboratory to everyday classrooms.

2.2 Material base for professional development: supporting teachers in noticing and supporting students in their basic conceptual needs

In order to achieve not only efficacy under laboratory conditions but also effectiveness of the intervention approach under field conditions, the material base had to be extended from 14 lessons to a two-year program that could be used by teachers with diverse mathematics education backgrounds and teaching experiences.

Thus, in the design research phase of Mastering Math (2011–2014), the pilot intervention program was elaborated into more comprehensive and robust teaching materials aimed at supporting teachers in noticing and fostering low-achieving students' conceptual understanding and basic competence (Selter, Prediger, Nührenbörger, & Hußmann, 2014).

The teaching materials are presented in 31 units covering the specified conceptual needs of low achievers, especially the place value system and meaning and representation of basic operations for natural and decimal numbers and fractions. All units are based on the three main design principles (Hußmann, Nührenbörger, Prediger, Selter, & Drücke-Noe, 2014) that form the Mastering Math teaching approach:

(1) Focusing on conceptual understanding. Rather than training fact retrieval or procedures, focusing on conceptual understanding allows students with mathematical difficulties to construct meanings for mathematical practices (Andersson, 2010; Boaler, 2002; Hiebert & Grouws, 2007; Prediger, Freesemann, Moser Opitz, & Hußmann, 2013). For example, understanding the place value system is crucial for flexibly calculating multi-digit numbers. Because a rule such as "multiplying by 10 always means shifting to the next digit" can only be understood by decomposing the numbers into hundreds (H), tens (Z for *Zehner* in German in Fig. 1), and ones (E for *Einer* in German) and connecting the digits to the visual representation.

(2) Ensuring adaptivity using diagnostic tasks. In order to support teachers to build on students' individual learning potentials and learning needs and help them deal with difficulties in diagnostic teaching (Bell, 1993; Sundermann & Selter, 2013), every teaching unit (consisting of 3–5 lessons) starts with a one-page formative assessment sheet with diagnostic tasks offering insights into student thinking. To analyze students'

answers with the intent of determining typical resources or misconceptions, the teacher's manual shows the most frequently occurring answers together with suggestions on how to interpret them and how to address them (Fig. 2).

(3) Promoting discourse. Conceptual understanding is best developed when students engage in mathematical discourses (Bell, 1993; Nührenbörger & Schwarzkopf, 2010). Thus, the teaching units have not been optimized for independent individual seatwork, but rather for small-group work with teachers' moderation on problems with multiple solutions and perspectives (such as in Fig. 1), encouraging students to communicate, explain, and argue.

Thus, the Mastering Math teaching approach consisted of the design principles together with the material for students and teachers. The materials were designed in order to enable teachers to identify those students with the most difficulties, to realize which basic conceptual needs they have, and to foster their nurturing of the students' conceptual needs, as shown in Fig. 1. The teacher's manual did not merely prescribe what teachers were expected to do but was also intended to be educative. The iterative design process for the materials took several design research cycles and ended with the scaled up implementation in the project architecture presented in the next section.

3 Multi-level project architecture for scaling up in Mastering Math NRW

A major design challenge for all scaling-up processes is to develop research-based implementation and professional development programs that help to bring innovations such as those presented in the Mastering Math program into classrooms. According to the literature base (see

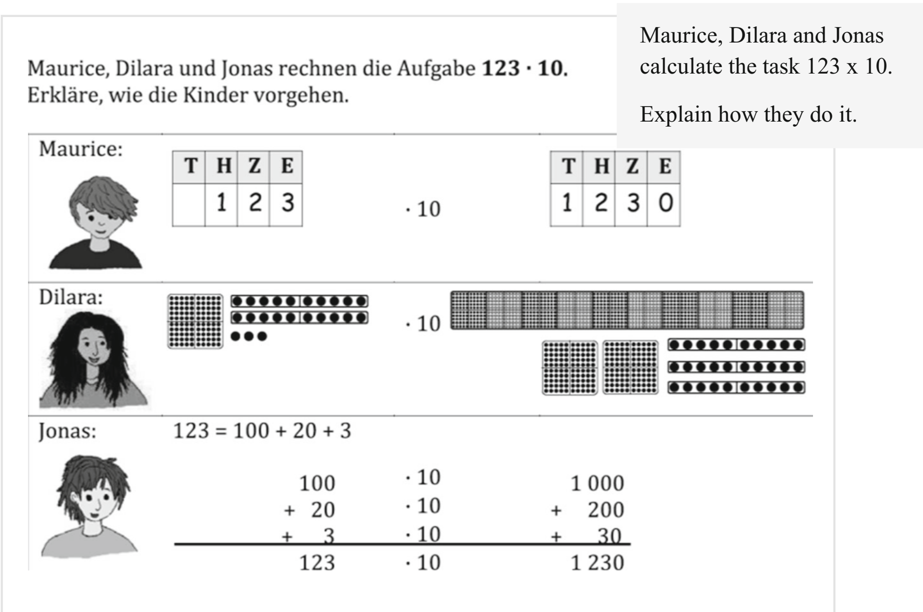

Fig. 1 Extract from the teaching material: connecting calculation to place value understanding (Selter et al., 2014, p. 59)

Item in the formative assessment

Trage in die Stellentafel ein und schreibe die Zahl auf.

	Stellentafel T	Stellentafel H	Stellentafel Z	Stellentafel E	Zahl
3 Tausender, 1 Zehner, 10 Einer	3	0	1	10	3 020
20 Hunderter, 4 Zehner		20	4	0	2040
6 Tausender, 2 Hunderter, 42 Zehner, 5 Einer	6	2	42	5	6625

Translation from German:
Fill in the place value table and write the number.
3 thousands, 1 ten, 10 ones
20 hundreds, 4 tens
6 thousands, 2 hundreds, 42 tens, 5 ones

Analytic support provided in the teacher's manual (translated from German)

Typical answers for (c)	Interpretation	Suggested starting point
62,425, e.g., "There are 6 T, 2 H, 42 Z, 5 E, thus 62,425"	Place value of digits not considered	Discuss place values of digits, connect to enactive representation in Task 2.1
6,607, e.g., "You have to pay attention with the hundreds and the tens"	Recognized the difficult aspect but does not yet know how to solve it.	Revise dealing with bundles of tens, see Task 2.2a
6,247, e.g., "Well, the 6 to T, the 2 to H, the 42 must be split, the 4 to Z (tens), and the 2 + 5 to E (ones)"	Decomposition of 42 tens into 4 tens and 2 ones	Decompose with enactive material, in Task 2.2b
6,620, e.g., "42 Z are 420"	Got the main point but overlooked the 5 ones	Just focus on forgotten ones, system is most probably understood

Fig. 2 Formative assessment and manual for interpreting the results (in the unit on place value understanding)

Section 1), we assume that this design challenge can be met using three implementation strategies: the material- and community-based implementation strategies and the systemic strategy, which involves addressing different levels of institutional contexts.

As this involves many decisions, this section presents the ways the design challenge was addressed in the project architecture of its implementation in North Rhine-Westphalia (referred to as Mastering Math NRW), for which this evaluation was conducted. For Mastering Math NRW, 40 comprehensive secondary schools (grades 5–13) and 120 teachers were involved and empowered to work with the developed teaching materials (material-based strategy of implementation).

In 2014, the implementation project started in grade 5 of the project schools; in 2015, it was extended to grade 6. The support provided by a project team that consisted of six researchers and one coordinator at TU Dortmund University who worked using the teaching approach and the materials to qualify nine network facilitators and to establish teacher communities in each school and networks across schools in order to engage the teachers in continuous cooperation (community-based strategy). The network facilitators moderated professional learning community meetings by following up on professional development using the materials and gradually handed over responsibility to the teachers. The entire complex project architecture with structures and support on five levels is summarized in Fig. 3.

As innovations can only be successful when embedded in institutional contexts (Prenzel et al., 2008), the project also involved the schools' principals and local and regional authorities and worked towards establishing networks of schools (systemic strategy).

The process of recruiting facilitators, schools, and teachers followed the sketched multi-level structure: First, decision-makers in the Ministry of Education had to be convinced. After their approval, the project leaders negotiated with the heads of several school districts in NRW who committed themselves to paying the network coordinators and recruiting schools willing to participate in the project over the long run.

The stakeholders in the school districts identified possible project schools and negotiated with the school principals in order to create the conditions necessary to implement the project (i.e., providing teacher resources and establishing reliable communication structures). Teachers who taught the Mastering Math groups in addition to the regular classes began to take part as volunteers at their schools when they felt confident with the intentions and the background philosophy of Mastering Math teaching approach. The Mastering Math project was then presented at a large conference to which the school principals and math teachers from the project schools were invited.

Levels	**Major tasks on each level**	**Support offered by Mastering Math team (6 researchers, 1 coordinator)**
Level of school districts	7 School districts • Recruit 40 schools to participate in the project • Pay 9 network facilitators to accompany teacher communities in the schools and to establish networks of schools in each district	• Negotiate with school districts to install the project • Qualify the network facilitators to facilitate the network communication and meetings (focus on main principles of material- and community-based strategies)
Level of schools as institutions	40 secondary comprehensive schools • Jointly commit themselves as being a Mastering Math school • Provide teaching resources for extra courses (1-3 lessons per week, depending on size of school) • Establish reliable communication time for teacher communities in schools (regularly visited by network facilitator)	• Negotiate with principals to install reliable structures for courses and teachers' community meetings • Sign written contracts with each school to establish reliable mutual expectations
Level of teacher communities in schools	40 teacher communities in schools • Regularly meet in the school, partly with the school's network facilitator • Discuss experiences of noticing and supporting students • Increasingly include the Mastering Math teaching approach in their teaching repertory (also in regular classrooms)	Qualify and accompany the 9 network facilitators to • accompany teacher communities and enhance cooperation • support inclusion of the teaching approach and teaching materials • support teachers in interpreting diagnostic tasks and in overcoming obstacles in the courses
Level of classrooms	90–120 teachers • Commit themselves as Mastering Math project members who work within the teaching approach and with the provided teaching materials • Teach one of the Mastering Math courses with 8-10 students each	• Provide teaching material to support teachers in noticing and fostering students' basic conceptual understanding
Level of students	343 low-achieving students in intervention • are selected out of 3,837 students in the 40 schools involved based on their basic needs • come to one lesson of a Mastering Math course per week in addition to regular classes (similarly in the control schools)	• Provide standardized screening tool to select students with basic conceptual needs • Pre- and post-test for assessing effectiveness

Fig. 3 Systemic multi-level structure of the project architecture in Mastering Math NRW (North Rhine-Westphalia)

In this project architecture, the nine network facilitators played a key role in navigating between the top-down, material-based strategy and the more bottom-up, community-based strategy. On the one hand, the network facilitators were qualified as experts in the Mastering Math teaching approach, the underlying materials for use in teaching, and the underlying teaching philosophy, including content-specific information on the mathematical background, students' ways of thinking, possible mistakes and alternative conceptions, and possible ways of working with students. The network facilitators were the ones who communicated these aspects to the teachers. On the other hand, they combined their input with activities designed to enhance the peer-to-peer communication and cooperation that had been installed.

In this respect, the facilitators were responsible for empowering the teachers' communities to work with increasing agency and independence. In the beginning, the network meetings across the schools of a district were mainly used for discussing the underlying teaching philosophy of the material, analyzing students' solutions, and planning the next course lessons based on reflecting previous experience. Increasingly, priorities shifted towards the development of strategies to incorporate the Mastering Math teaching approach into regular classroom teaching and in the long run into the schools' structures. The latter is important for sustainability after the end of the implementation phase.

Summing up, we combined the four types of support for implementing innovation described by Cobb and Jackson (2012):

- **New positions**. An example of this is the addition of one person per district/school responsible for organizing the professional development in mathematics education. In our project, one network coordinator and nine network facilitators were appointed.
- **Learning events**. This includes professional development, either intentional or unintentional, and on-going or discrete events. In our project, this was realized by regular professional development meetings in the schools, in the network, or across all nine networks that offered input on the teaching materials and their backgrounds and by peer-to-peer exchange in the teachers' communities in each school and the network meetings of all schools involved in one district.
- **New organizational routines**. These can include regular meetings of the principals of schools and all mathematics teachers. In our project, this was realized by installing the teacher communities in each school. These groups met regularly for joint work in fixed time slots, sometimes visited by the schools' network facilitator, and they established additional courses for students at risk.
- **New tools**. These could consist of classroom materials, materials for professional development, or revised curriculum frameworks. In our project, this aspect was considered very crucial, so complete teaching material for two years of additional courses was created.

4 Methods for investigating the effects on different levels

The description of the project architecture in Section 3 has high practical relevance, which can translate into theoretical relevance if the project can show its effectiveness. For this purpose, we present some of the methods applied for the external evaluation conducted by IPN Kiel and data collected in the first two years of the implementation project. The results are presented in Section 5.

4.1 Methods of data gathering

4.1.1 Measures and instruments

The mixed-methods evaluation captured conditions and effects using several measures and instruments. The choice of measures takes into account the different levels (see Fig. 3) and different domains of knowledge (project management, process of teaching and professional development, effects on teaching, and teachers and students).

- On the student level, the learning gains of the intervention were measured using BASIS MATH G4+-5 (Moser Opitz et al., 2016) to perform standardized pre- and post-tests at the beginning of grade 5 (September 2015) and beginning of grade 6 (September 2016). The focus was on basic conceptual needs in arithmetic (place value understanding, meanings of operations, and word problems) and basic arithmetic skills. The standardized cutoff for identifying low-achieving students at risk is a score of 29 out of 60 (and 31 for the post-test).
- On the teacher level, an online questionnaire was administered using the free service platform LimeSurvey. The instrument was adapted from other teacher professionalization projects (Fischer & Rieck, 2014) and used four Likert-type scales (with the anchors being *fully agree*, *partially agree*, *partially disagree*, and *fully disagree*). In general, self-reported data must be taken with care as they do not capture objective facts but the subjective views of the participants. We interpreted the subjective views as indications of teachers' acceptance of Mastering Math, as acceptance is always a precondition for incorporating innovations in regular professional practice (Rogers, 2003).
- On the teacher community and school level, the protocols of meetings on all levels were collected (e.g., facilitators' meeting the project team at the university).

4.1.2 Samples

The comparative intervention study involved 40 schools in the implementation project and 21 control schools that were offered integration into the project's work two years later. The students who participated in the intervention and control groups (about 15% of all students) were the low-achieving students who participated in additional remediating courses in classrooms using either the Mastering Math approach or other programs at the control school (Fischer, Schöber, Döring, & Köller, 2017). Table 1 shows that both subsamples were comparable in terms of the immigrant statuses of the students (slightly higher in the control group) and pre-test scores.

The online questionnaire was administered two months after the second year of the assessed intervention started, only to the teachers in the Mastering Math schools and not to those in the control schools. Sixty-three teachers answered, and 59 answered all questions. The teachers were diverse in their job experience (56% had taught less than 10 years, 31% 11–20 years, and 14% longer than 20 years). Forty-four percent were active in the project from 6 to 14 months, and a majority (56%) joined the project two months before (Fischer, Rieck, Döring, & Köller, 2017).

4.2 Methods of data analysis

The test and questionnaire data were analyzed quantitatively using SPSS 19. For the questionnaires, only frequencies of agreement on selected items are reported here. For comparing

Table 1 Characterizing the involved schools and intervention subsamples

	Whole samples		Subsamples in intervention study	
	All students at Mastering Math schools	All students at control schools	Intervention group: low achievers in Mastering Math courses	Control group: low achievers in other additional courses in control schools
Number of schools/number of classes	40/165	21/83	See left	See left
Number of students (N)	3837	1898	592	389
Students with immigrant status	29.3%	34.7%	61.7%	52.2%
Pre-test score mean (SD)	34.81 (11.80)	33.93 (12.24)	19.12 (6.52)	19.65 (6.67)
Low-achieving students (pre-test below cutoff 29)	31%	35%	100%	100%

students' learning gains of intervention and control group (Research Question RQ2), a repeated measure analysis of variance (one-way ANOVA) was conducted on the scores of the pre- and post-test. Group and time were the main factors of the ANOVA, and group by time was the interaction factor for the analysis. Within these analyses of variance, the inter-group effect sizes were determined using partial eta squared.

More complex answers to RQ3 were provided by insights into the qualitative analysis of protocols, mainly from the facilitators' meetings with the research team. For this purpose, a content analysis was conducted to identify the most crucial chances and challenges and their interplay on different levels.

5 Results of the quantitative and qualitative analyses

5.1 Effects on the teacher and school level (RQ1)

This section addresses RQ1: What are the effects of the implementation and professional development program on teachers' perceptions of materials and of their cooperation in professional learning communities? The questionnaires offer—among other things—interesting insights into the teachers' self-reports on the perceived use of the material provided by Mastering Math and into challenges in the field of cooperation in the professional communities of the teachers (see Table 2).

Teachers' self-reported effects of the material-based strategy With respect to the offered teaching materials, 76.3% of the teachers felt supported by the Mastering Math material and ranked this support highest, while 16.9% partially agreed to this item. However, concrete hints on how to deal with the material in practice were important for 41.3% (33.3%) and another 30.5% considered the in-service training helpful. The facilitators played an important role in the implementation of both the project's concept and structures and the material. Sixty-five point one percent of the teachers felt supported by the facilitators' input on the content level (17.5% partially agreed), for 62% the explanation of the material was crucial (19% partially agreed), and 41.3% highly estimated the ideas about how the material can be used in classroom practice (33.3% partially agreed). Forty-four percent felt it was helpful to receive constructive feedback in the domain of diagnosis and support of students (27.1% partially agreed).

Table 2 Teachers' perceptions of materials and cooperation

Item	Percentages of full/partially agreement to the items among the responding teachers
Perceived support by material and additional advice	
Support by project material	76.3% / 16.9%
Support by concrete hints for working with the material	41.3% / 33.3%
Support by in-service training	30.5% / 30.5%
Support by facilitators' content input	65.1% / 17.5%
Support by facilitators' explanations of project material	61.9% / 19.0%
Perceived support by the learning community	
Support by common aims in the group	68.5% / 18.5%
Support by exchange with other teachers	47.5% / 35.8%
Support by exchange of material	51.9% / 20.4%
Support by task sharing	55.6% / 22.2%
Support by group work being more effective than individual work	38.9% / 25.9%

Teachers' self-reported effects of the community-based strategy Only 47.5% of the respondents found the exchange with other teachers very supportive, while 35.6% partially agreed to this item (see Table 1). Thus, implementing professional communities of teachers requires special investment in establishing such groups. The difficulties can be read off by the following data: Usually, a cooperating unit consists of at least three people. Smaller units cannot be considered a group. In 2015, 35.5% of the teachers worked on the school level with one other partner, 4.8% were alone, 44.4% were active in groups of three to five, and 14.3% in groups of six to ten persons. We concluded from this finding that only 58.7% were already engaged in professional collaboration; for the other group, we saw urgent need for intervention.

On the content level, cooperation includes common aims and common actions referring to those aims. Teachers reflect on the classroom practice, identify the needs of the students, and find the suitable material. They alter existing material, develop if necessary additional material, and give each other feedback. Teaching is a constructive process, and the implementation of a new teaching practice requires reflecting, restructuring, and reconstructing the present practice. As far as the quality of cooperation is concerned, 68.5% of the respondents self-reported common aims that were in line with the project's concept. These teachers jointly reflected on their teaching. 51.9% altered the teaching material and reflected on the project work, while 31.5% developed additional teaching material. With respect to the cooperation practices, 55.6% reported successful task sharing. Thirty-eight point nine percent perceived group work as more effective than individual work (25.9% partially agreed).

These findings correspond to a general problem in German school culture: There is no tradition of professional cooperation. Even for teachers who volunteered to participate in Mastering Math, developing professional collaboration required more support.

Reported challenges on the schools as institution level Installing teachers' professional learning communities is also a task for schools as institutions, especially for principals. The questionnaire showed that starting the project with only one or two colleagues in a school hindered establishing teacher groups for nearly 40% of the teachers. The majority of the respondents reported a lack of support by the school principals and a lack of interest of the teachers who were not involved. Only 35% of the respondents reported that the principals sufficiently supported the implementation process. Thirty-eight percent stated that their principal was interested in Mastering Math, and the same percentage saw their principal implementing it. Only 9.5% heard that their principal had informed the teachers not involved that the project was occurring. These formative insights into obstacles in the program at the end of the first year were treated by intensifying communication with the principals (see Section 5.3).

5.2 Effects on students' learning gains (RQ2)

The achievement scores and ANOVA results for both subsamples for RQ2 (What are the effects of the research-based implementation and professional development program on students' learning gains compared to a control group?) are shown in Table 3. The intervention and control groups show impressive increases in the test scores, with effect sizes of $d > .80$. In the ANOVA, the significance of the main factor, time, with $F_{(\text{time})} = 1246.6$ and an effect size of $\eta^2 = .560$, suggests that as expected, the low-achieving students in both groups improved their basic conceptual understanding and skills during one year of schooling.

The significant interaction effect, with $F_{(\text{group} \times \text{time})} = 31.26$, $p < .001$, suggests that the intervention group had significantly higher learning gains than the control group from pre- to post-test. The effect size of $\eta^2 = .031$ is considerable for field studies (Cheung & Slavin, 2016). The post hoc test does not show a pairwise difference between both groups, showing empirical evidence for the effectiveness of the program on the student level.

Even if the effectiveness is shown by comparing learning gains, other parts of the analysis provided unexpected results: For example, the length of students' participation in the intervention did not seem to influence learning gains, nor did the teachers' self-identification as a Mastering Math teacher. For these reasons, deeper insights into the processes are required.

5.3 Challenges and advantageous conditions on other levels (RQ3)

Although the data analysis presented here is focused on the teacher, the teacher community, and the school level, it is worth complementing it by snapshots from other levels that we drew

Table 3 Comparison of leaning gains between intervention and control group

	Scores in pre-test m (SD)	Score in post-test m (SD)	Gains in average scores from pre- to post-test	Intra-group effect size d from pre- to post-test
Intervention group: low-achieving students in Mastering Math courses	19.12 (6.52)	29.44 (10.08)	$\Delta = 10.32$	$d = 1.24$
Control group: low-achieving students in other additional courses in control schools	19.65 (6.67)	27.15 (10.10)	$\Delta = 7.5$	$d = 0.89$
Results of ANOVA: inter-group and time effect from pre- to post-test	$F_{(\text{time})} = 1246.6$, $p < .001$, $\eta^2 = .560$; $F_{(\text{group})} = 3.18$, $p = .08$ (n.s.), $\eta^2 = .003$; $F_{(\text{group} \times \text{time})} = 31.26$, $p < .001$, $\eta^2 = .031$			

from the minutes of the facilitators' meetings within schools, across schools in network meetings, and across networks in meetings with the project team at the university. These data are used to deal with RQ3: What are possible challenges and advantageous conditions for the program on the different levels?

The necessity of including the school district level already became visible before the project started: Whereas some district authorities were very supportive in recruiting schools and facilitators, others had more difficulties. During the two and a half years of project work, the most supportive district was the one with the best functioning schools. For the principals, the relevance assigned by the district authority to an implementation project was a crucial determinant for the relevance assigned within the district's school. This had immediate impact on other levels.

The schools as institutions level was crucial for installing the Mastering Math courses as additional courses in the school's time tables and for installing reliable meeting times for the school groups functioning as professional learning communities. The more the principals were committed to the project, the easier were the teachers' institutional contexts. When problems became visible in the principals' commitments (see Section 5.2), all 40 principals were invited again to meet the project team at the university and discuss challenges and means of possible support. For some schools, the principals' commitment could be strengthened, which led to much better working conditions in the proceeding months.

The teacher communities in school level proved to be very important in some schools where teachers started to engage in intense communication and cooperation. In other schools, however, the large size of the groups turned out to be an obstacle for experimenting with innovations, so that teachers split up in order to gain the freedom to experiment. In other schools, the lack of adequate group size mentioned above (just one or two teachers involved) turned out to be an advantage to getting the work started, as these isolated teachers were not slowed down by doubting colleagues and found their professional learning community in the network meeting across schools in the district. However, in order to establish the courses in the schools in the long run, these teachers working in isolation will need to find allied colleagues.

On the classroom level, establishing extra courses with only 8–10 students proved to be of major importance for the teachers' professionalization pathways. Although difficult to establish in some schools (cf. school level), these small courses allowed teachers a much more intense communication with their low-achieving students. These were important conditions for increasing the sensitivity for students' conceptual needs. In the third year, the teachers are now ready to integrate the Mastering Math teaching approach and material into their regular classrooms, which have 25–30 students. For the first two years, limiting the size to small groups was also an excellent opportunity to extend the teachers' repertoires of teaching styles to be less directive and diagnostic.

The existence of 31 complete teaching units with carefully planned learning trajectories and diagnostic tasks proved to be a successful curricular orientation. Teachers' interest in formative assessment increased only when experiencing the added value of the diagnostic information for their teaching. Once they engaged in discussions about students' different ideas and conceptions, they developed an appreciation for the need of formative assessment. The material base on the teaching unit level rather than single tasks provided a good base for offering experiences of success. On the one hand, this allowed them to become familiar with the principles of promoting discourse and focusing on conceptual understanding, which are crucial for their professional development. On the other hand, the project team will encourage them in the future to select the most important parts in order to overcome the time problems that occur when every single task is used. Hence, a material base is very helpful, but teachers need to keep the agency in selecting the parts that are most relevant for their students.

6 Discussion and conclusion

Should a community-based implementation or material-based implementation strategy be used? Many researchers emphasize the need to combine them for achieving the scaling up of implementation projects (Lachance & Confrey, 2003; Llinares & Krainer, 2006). However, there is still a lack of evaluated projects that realize the recommended combination of strategies. Mastering Math NRW can serve as a case that can show the complexities of combining the strategies and addressing different levels (systemic strategy) and provide empirical evidence for its effectiveness under field conditions.

6.1 Limitations of the study

Of course, the specific ways of combining the three strategies are tightly connected to the specific conditions of the presented project and its very specific focus on low-achieving students in grades 5 and 6. Additionally, the research revealed complexities that raise challenges that need to be overcome in the years following the project. The project also has methodological limitations: Because schools were not randomly assigned to the intervention and control groups but rather volunteered for the program, no classroom observations were conducted in order to account for teachers' practices, and the number of involved teachers was limited. The intergroup effect size of $\eta^2 = .031$ for comparing students' learning gains is an even greater limitation, as the study was only organized as a quasi-experimental study. However, considering the size of the sample and the difficulty of the field conditions, with teachers starting immediately, the effect size is still considerable according to the meta-study of Cheung and Slavin (2016).

However, the project can already provide some substantial answers to the design challenge (DC) and research questions (RQ1–3) that allow refining the material- and community-based strategies, as we do below.

6.2 Material-based elements of the implementation strategy and their effects

Implementing an innovative teaching approach can start from laboratory products for which efficacy is shown in randomized or quasi-randomized controlled trials. However, before widely implementing it in schools, further years of design research may be required to transform the laboratory intervention program into more robust teaching materials that are applicable for many teachers (Burkhardt, 2006). As the current study has shown, teaching materials can form a material base for scaling up when following these three characteristics:

- The material supports teachers to pursue critical quality criteria, in our case (1) focusing on conceptual understanding, (2) ensuring adaptivity using diagnostic tasks, and (3) promoting discourse (Hußmann et al., 2014) that are in line with general demands in effective mathematics education approaches (Hiebert & Grouws, 2007).
- The material base encompasses not only isolated tasks, but 31 complete teaching units for two years (as recommended by Gravemeijer et al., 2016), each with well-planned learning trajectories and corresponding diagnostic tasks and a teacher's manual.

- The material is not only delivered to the schools, but implemented within a multi-level support structure (intertwinement with a systemic strategy).

The reported investigation of the teachers' questionnaires showed that the offered material base provided important support for the teachers, as it offered a clear curricular orientation about what to do and sensitized them to students' basic conceptual needs that had previously been neglected (RQ1).

6.3 Community-based elements of the implementation strategy

The material-based implementation strategy should be combined with a community-based strategy to offer teachers a communicative space to develop their own ways of realizing the teaching approaches based on their own school contexts. The formative evaluation showed that teachers need one year of facilitation, which can help realize the community-based strategy along the following lines:

- Creating long-term professional learning communities in each school ("school groups") with regular joint meeting times, jointly declared aims, and binding communicative structures. Too-small school groups should be combined across several schools.
- Organizing the communities' work around the materials by discussing lesson plans and reflecting experiences, especially with respect to examining and analyzing children's thinking. This principle harmonizes the top-down and bottom-up approaches of the material-based and community-based strategies and is a crucial contribution to the existing literature.
- Including the other institutional levels (principals and district levels) as far as possible in order to give a supportive background structure. This point needs to be strengthened in the next project.

The qualitative research on the oral feedback, observations, and questionnaires showed that it takes time to install professional learning communities with adequate working conditions within the schools (RQ1), but the installation was heavily strengthened by the existing material as a common base. For many German schools, meeting regularly with colleagues is a real innovation for which the organizational structures (such as a joint hour in the weekly time table) first have to be established (RQ3). In spite of these difficulties and complexities, the students' learning gains were already significant in the first year (RQ2).

6.4 Systemic elements of the implementation strategy

The facilitators played a key role in combining the material-based strategy with the community-based strategies as they helped the teachers' communities within and across schools to adapt the teaching approaches in the teaching materials to their own school contexts with specific students and conditions (RQ3).

Our overall experience in the project showed that the material-based and community-based strategies can be combined effectively and have an effect even on the level of students' learning gains. In addition, it is of crucial importance not to neglect the systemic strategy. Further research will need to disentangle differential effects of different design elements rather than test all together.

Acknowledgements We thank the project team in Dortmund (Sabrina Lübke, Stephan Hußmann, Corinna Mosandl, Marcus Nührenbörger, Birte Pöhler, Gerd Seifert, and Lara Sprenger), the evaluation team in Kiel (Franziska Trepke, Brigitte Döring, and Olaf Köller), and all the involved facilitators, schools, and teachers.

Funding information The research was funded by the German Telekom Foundation and is conducted within the DZLM research program (Deutsches Zentrum für Lehrerbildung Mathematik: German Center for Mathematics Teacher Education).

References

Adler, J., & Jaworski, B. (2009). Public writing in the field of mathematics teacher education. In R. Even & D. Ball (Eds.), *The professional education and development of teachers of mathematics* (pp. 249–254). New York: Springer.

Andersson, U. (2010). Skill development in different components of arithmetic and basic cognitive functions. *Journal of Educational Psychology, 102*(1), 115–134.

Bell, A. (1993). Some experiments in diagnostic teaching. *Educational Studies in Mathematics, 24*(1), 115–137.

Boaler, J. (2002). *Experiencing school mathematics*. Mahwah: Lawrence Erlbaum.

Bonsen, M., & Rolff, H.-G. (2006). Professionelle Lerngemeinschaften von Lehrerinnen und Lehrern [Professional learning communities of teachers]. *Zeitschrift für Pädagogik, 52*(2), 167–185.

Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. Q. (2010). *Organizing schools for improvement*. Chicago: University of Chicago Press.

Burkhardt, H. (2006). From design research to large scale impact. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 121–150). London: Routledge.

Burkhardt, H., & Schoenfeld, A. (2003). Improving educational research. *Educational Researcher, 32*(9), 3–14.

Cheung, A. C., & Slavin, R. E. (2016). How methodological features affect effect sizes in education. *Educational Researcher, 45*(5), 283–292.

Cobb, P., & Jackson, K. (2012). Analyzing educational policies: A learning design perspective. *The Journal of the Learning Sciences, 21*, 487–521.

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher, 32*(6), 3–12.

Darling-Hammond, L. (1997). Restructuring schools for student success. In A. H. Halsey, H. Lauder, P. Brown, & A. Stuart Wells (Eds.), *Education—Culture—Economy, and Society* (pp. 332–353). Oxford: Oxford University Press.

Darling-Hammond, L., & Richardson, N. (2009). Teacher learning: What matters? *Educational Leadership, 66*(5), 46–53.

Euler, D., & Sloane, P. (1998). Implementation als Problem der Modellversuchsforschung [Implementation as a problem of field research]. *Unterrichtswissenschaft, 26*(4), 312–326.

Fischer, C., & Rieck, K. (2014). Improving teaching in science and mathematics. In R. E. Slavin (Ed.), *Classroom management and assessment. Proven programs in education* (pp. 110–115). Corwin: Thousand Oaks.

Fischer, C., Rieck, K., Döring, B., & Köller, O. (2017). Externe Evaluation von „Mathe sicher können“. *Ergebnisse der Gesamtbefragung der Lehrkräfte* [External evaluation of “Mastering Math”. Results of the teacher survey]. Kiel: IPN. http://mathe-sicher-koennen.dzlm.de

Fischer, C., Schöber, C., Döring, B., & Köller, O. (2017). *Externe Evaluation von “Mathe sicher können”. Ergebnisse der Testung der Lernenden* [External evaluation of “Mastering Math”. Results of the student assessment]. Kiel: IPN. http://mathe-sicher-koennen.dzlm.de

Gräsel, C., Fußangel, K., & Pröbstel, C. (2006). Die Anregung von Lehrkräften zur Kooperation—eine Aufgabe für Sisyphos? [Initiating teachers’ cooperation. A task for Sisyphos?]. *Zeitschrift für Pädagogik, 52*(2), 205–219.

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.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students’ learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Charlotte: Information Age.

Hußmann, S., Nührenbörger, M., Prediger, C., Selter, C., & Drücke-Noe, C. (2014). Schwierigkeiten in Mathematik begegnen [Overcoming difficulties in mathematics]. *Praxis der Mathematik, 56*(56), 2–8.

Krainer, K. (2008). Individuals, teams, communities and networks: Participants and ways of participation in mathematics teacher education. In K. Krainer & T. Wood (Eds.), *International handbook of mathematics teacher education* (Vol. 3, pp. 1–10). Rotterdam: Sense.

Kullmann, H. (2012). Lesson Study—eine konsequente Form unterrichtsbezogener Lehrerkooperation. In S. G. Huber & F. Ahlgrimm (Eds.), *Kooperation. Aktuelle Forschung zur Kooperation in und zwischen Schulen sowie mit anderen Partnern* (pp. 69–88). Waxmann: Münster.

Lachance, A., & Confrey, J. (2003). Interconnecting content and community: A qualitative study of secondary mathematics teachers. *Journal of Mathematics Teacher Education, 6*(2), 107–137.

Llinares, S., & Krainer, K. (2006). Professional aspects of teaching mathematics. In A. Gutiérrez & P. Boero (Eds.), *Handbook of research on the psychology of Mathematics Education* (pp. 429–459). Rotterdam: Sense.

Lomos, C., Hofman, R. H., & Bosker, R. J. (2011). Professional communities and student achievement—A meta-analysis. *School Effectiveness and School Improvement, 22*(2), 121–148.

Maaß, K., & Artigue, M. (2013). Implementation of inquiry-based learning in day-to-day teaching: A synthesis. *ZDM, 45*(6), 779–795.

Maccini, P., Mulcahy, C. A., & Wilson, M. G. (2007). A follow-up of mathematics interventions for secondary students with learning disabilities. *Learning Disabilities Research & Practice, 22*(1), 58–74.

McDuffie, A. M., & Mather, M. (2006). Reification of instructional materials as part of the process of developing problem-based practices in mathematics education. *Teachers and Teaching: Theory and Practice, 12*, 435–459.

Moser Opitz, E. (2007). *Rechenschwäche/Dyskalkulie* [Mathematical learning difficulties/Discalculia]. Bern: Haupt.

Moser Opitz, E., Freesemann, O., Grob, U., & Prediger, S. (2016). *BASIS-MATH-G 4+-5. Gruppentest zur Basisdiagnostik Mathematik* [Group test for basic assessment mathematics]. Bern: Hogrefe.

Moser Opitz, E., Freesemann, O., Prediger, S., Grob, U., Matull, I., & Hußmann, S. (2017). Remediation for students with mathematics difficulties. *Journal of Learning Disabilities, 50*(6), 724–736.

Nührenbörger, M., & Schwarzkopf, R. (2010). Die Entwicklung mathematischen Wissens in sozial-interaktiven Kontexten [the development of mathematical knowledge in social-interactive contexts]. In C. Böttinger, K. Bräuning, M. Nührenbörger, R. Schwarzkopf, & E. Söbbeke (Eds.), *Mathematik im Denken der Kinder* (pp. 73–81). Seelze: Klett-Kallmeyer.

Prediger, S., Freesemann, O., Moser Opitz, E., & Hußmann, S. (2013). Unverzichtbare Verstehensgrundlagen statt kurzfristige Reparatur [Indispensable basic needs rather than short-term repair]. *Praxis der Mathematik, 55*(51), 12–17.

Prenzel, M., Friedrich, A., & Stadler, M. (2008). *Von Sinus lernen. Wie Unterrichtsentwicklung gelingt* [Learning from the model project Sinus. How classroom development can succeed]. Seelze: Kallmeyer.

Prenzel, M., Sälzer, C., Klieme, E., & Köller, O. (Eds.). (2013). *PISA 2012*. Münster: Waxmann.

Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research, 75*(2), 211–246.

Roesken-Winter, B., Hoyles, C., & Bloemeke, S. (2015). Evidence-based CPD: Scaling up sustainable interventions. *ZDM Mathematics Education, 47*(1), 1–12.

Rogers, E. M. (2003). *Diffusion of innovations*. New York: The Free Press.

Schellenbach–Zell, J., Rürup, M., Fussangel, K., & Gräsel, C. (2008). Bedingungen erfolgreichen Transfers am Beispiel von Chemie im Kontext [Conditions of successful transfer for the example of Chemistry in Context]. In R. Demuth, C. Gräsel, B. Ralle, & I. Parchmann (Eds.), *Chemie im Kontext* (pp. 81–121). Münster: Waxmann.

Selter, C., & Bonsen, M. (2018). Konzeptionelles und Beispiele aus dem Projekt PIKAS [Ideas and examples from the project PIKAS]. In R. Biehler, T. Lange, T. Leuders, B. Roesken-Winter, P. Scherer, & C. Selter (Eds.), *Mathematikfortbildungen professionalisieren* (pp. 143–164). Heidelberg: Springer.

Selter, C., Prediger, S., Nührenbörger, M., & Hußmann, S. (Eds.). (2014). *Mathe sicher können—Natürliche Zahlen. Diagnose- und Förderkonzept* [Mastering Math—Natural numbers. Material for formative assessment and intervention]. Berlin: Cornelsen.

Selter, C., Wessel, J., Walther, G., & Wendt, H. (2012). Mathematische Kompetenzen im internationalen Vergleich. Testkonzeption und Ergebnisse. In W. Bos, H. Wendt, O. Köller, & C. Selter (Eds.), *Mathematische und naturwissenschaftliche Kompetenz von Grundschulkindern* (pp. 69–122). Münster: Waxmann.

Sundermann, B., & Selter, C. (2013). *Beurteilen und Fördern im Mathematikunterricht* [Assessing and fostering in mathematics classrooms]. Berlin: Cornelsen.

Swan, M. (2007). The impact of task-based professional development on teachers' practices and beliefs. *Journal of Mathematics Teacher Education, 10*(4–6), 217–237.

Takahashi, A., & Yoshida, M. (2004). Lesson-study communities. *Teaching Children Mathematics, 10*(9), 436–437.