ORIGINALARBEIT/ORIGINAL ARTICLE

Orchestrating productive classroom talk in Swiss second grade mathematics classrooms

Christian Rüede · Christine Streit · Sog Yee Mok · Rahel Laubscher

Abstract Productive classroom talks provide a variety of learning opportunities including the collaborative construction of mathematical meaning. To support the orchestration of productive classroom talk, teachers can use specific conversational strategies called "talk moves". Positive relations between the number of productive talk moves, discourse quality, and learning success have been demonstrated, but mostly at the secondary school level and in English-speaking countries. There are very few findings from elementary school in German-speaking countries. The present study makes the first step towards closing this gap by analyzing classroom talk in mathematics second grade classes in Switzerland ($n = 22$). The dataset includes recordings of mathematics lessons as well as students' pretest and posttest performance in mathematics. The results show that teachers varied in their use of productive talk moves and that some talk moves such as reasoning moves were preferred, depending on the teacher. In addition, the number of productive talk moves was positively related to both the number of justifications students provided and their learning success. Implications for the use of productive classroom talk in elementary school mathematics lessons are discussed.

Keywords Productive classroom talk · Student justification · Elementary school

✉ Christian Rüede · Christine Streit · Rahel Laubscher
Teacher School, FHNW University of Applied Sciences and Arts Northwestern Switzerland,
Muttenz, Switzerland
E-Mail: christian.rueede@fhnw.ch

Sog Yee Mok
University of Teacher Education Grisons, Chur, Switzerland

Springer

Orchestrierung produktiver Unterrichtsgespräche im Schweizer Mathematikunterricht der zweiten Primarstufe

Zusammenfassung Produktive Unterrichtsgespräche ermöglichen vielfältige Lerngelegenheiten, u. a. durch die gemeinsame Konstruktion von mathematischer Bedeutung. Zur Unterstützung der Führung produktiver Unterrichtsgespräche können Lehrpersonen spezifische Gesprächsstrategien, sogenannte produktive „talk moves" verwenden. In jüngster Zeit konnten positive Zusammenhänge zwischen der Anzahl produktiver „talk moves", Diskursqualität und Lernerfolg nachgewiesen werden, allerdings mehrheitlich in der Sekundarstufe und im englischsprachigen Raum. Für die Grundschule im deutschsprachigen Raum liegen dazu fast keine Befunde vor. Die vorliegende Studie macht erste Schritte zur Schliessung dieser Lücke: Wir analysierten Unterrichtsgespräche im Mathematikunterricht der zweiten Jahrgangsstufe in der Schweiz ($n = 22$). Der Datensatz umfasst Audiografien von Mathematiklektionen sowie Prä- und Posttestleistungen der Schülerinnen und Schüler in Mathematik. Die Resultate zeigen, dass die Lehrpersonen produktive „talk moves" unterschiedlich einsetzten und dass gewisse „talk moves" (z.B. Aufforderungen zur Begründung) bevorzugt wurden, abhängig von der Lehrperson. Zudem hängte die Anzahl produktiver „talk moves" sowohl mit der Anzahl von Schülerbegründungen als auch mit dem Lernerfolg positiv zusammen. Wir diskutieren Konsequenzen, die für den Einsatz produktiver Unterrichtsgespräche im Mathematikunterricht der Grundschule folgen.

Schlüsselwörter Produktives Unterrichtsgespräch · Schülerbegründung · Grundschule

1 Introduction

Classroom talks in mathematics are characterized by a specific type of communication called mathematical discourse. This *mathematical discourse* includes special words (e.g., "addition", "equals"), visual mediators (e.g., symbols, materials), routines (e.g., solution strategies), and relations between them, that are taken to be true (Sfard 2008). From a situated perspective, it is widely accepted in mathematics education that learning mathematics can be conceptualized as participation in *mathematical discourse practices* (Moschkovich 2015; Sfard 2008). These practices include ways of defining, describing, explaining, reasoning, justifying, and proving (Barwell 2020; Moschkovich 2007). Curricular recommendations regarding these practices can be found in the United States: Kilpatrick et al. (2001; adaptive reasoning); Germany: Kultusministerkonferenz (2022; mathematical communicating and arguing); and Switzerland: Deutschschweizer Erziehungsdirektorenkonferen (2016; exploring and arguing).

Participating in classroom talks provide opportunities for students to participate in mathematical discourse practice. However, the research literature provides ample evidence that while teacher-guided classroom talks are very commonly used in everyday teaching, their effectiveness for student learning is often assessed as

low (for an overview, see Howe and Abedin 2013). For example, the widespread dominance of the initiation-reply-evaluation-pattern (IRE-pattern, Mehan 1979) is critically discussed because it is often structured in small steps and its cognitive demand is mostly low. To counter such patterns in teacher-student verbal interactions, characteristics of classroom talks have recently been identified to support robust learning for each student (for a review, see Howe et al. 2019). Accordingly, teachers should encourage students to say more, to link their own answers with the answers of other students, and to reason and think together. So-called *productive classroom talks* have these characteristics (Alexander 2020; Michaels and O'Connor 2015). Studies at the secondary school level confirm that productive classroom talks positively affect the quality of student contributions and student performance (for a review, see Resnick et al. 2010).

Classroom talk is increasingly of interest for educators at the elementary school[1] level. Many students do not start learning mathematics systematically until they enter school. It is therefore all the more important that they develop their own personally meaningful ways of mathematical knowing during elementary school. Productive classroom talks can initiate and support the interactive constitution of mathematical meaning (Schütte et al. 2021; Yackel and Cobb 1996). There are different possibilities to integrate such classroom talks in mathematics lessons. For example, Häsel-Weide and Nührenbörger (2021) propose segmenting the lesson in different phases (see Stein et al. 2008). In the launch phase, a classroom talk can be used to initiate students' exploratory activities, in the reflection phase, it can be used to analyze and systematize their learning experiences. Studies focusing on mathematical discourse in elementary school have so far mainly investigated mathematical discourse in the explore phase, that is, discussions between two students or in small groups of students (Gellert 2014; Krummheuer 2011; Nührenbörger and Steinbring 2009). Few studies have analyzed mathematical discourse in the launch and reflection phase, that is, in classroom talk with the whole class.

The present study focused on the launch and reflection phase. In these phases, there are teacher-guided classroom talks. We investigated how productively teachers interact with students in this talk in classes with second grade students during mathematics lessons in Switzerland. We analyzed teachers' and students' verbal contributions during classroom talk, student performance, and the relationships between these variables. Our data included one documented mathematics lesson per class ($n = 22$) and each student's pretest and posttest performance on a mathematics test (beginning and end of school year).

---

[1] We use the terms "elementary school" and "elementary school level" to denote the first four grade levels. In Switzerland, these are the first four "Primarschulklassen" which in Germany are usually referred to as "Grundschule". We use "secondary school level" to denote grades five, six, seven, and higher.

2 Theoretical Background

To facilitate students' participation in mathematical discourse practices, specific learning opportunities should be provided in mathematics teaching. The present study focuses on such learning opportunities in elementary school classroom talks.

2.1 Orchestrating Productive Classroom Talk from an Educational Perspective

Classroom talk is understood as a prototypical form of teaching "in which teacher and students interact publicly, with the intent that all students participate (at least by listening)" (Hiebert et al. 2003, p. 53). It is distinct from individual, partner, or small group work. Educational research on classroom talk suggests its essential importance for teaching (Resnick et al. 2015). However, learning in a classroom talk depends on how the talk is guided. One type of classroom talk that aims to encourage students to collaboratively negotiate and construct mathematical meaning (e.g., of solution strategies, symbol representations) by participating in mathematical discourse practices is called *productive classroom talk*.

2.1.1 Productive Classroom Talk and Productive Talk Moves

Different but related generic frameworks of classroom talk have been developed under the umbrella of "productive classroom talk" (Alexander 2020; Michaels and

Table 1 Productive talk moves, adapted from O'Connor et al. (2017) and Chen et al. (2020)

Goal 1 (Elaborating). Encourage students to share, expand, or clarify their own ideas	
a	Ask a student to provide new information about the topic he/she addressed "Can you say more about that?" (Say more)
b	Repeat or rephrase some or all of what a student has said and provide the opportunity for him/her to verify it "So, are you saying ...?" (Revoice)
c	Indicate that students should pause and think "Turn and talk to the person next to you about that" (Time to think)
d	There is a pause lasting five and more seconds (Wait time)
Goal 2 (Listening). Encourage students to listen to one another	
a	"Who can repeat what (s)he just said?" (Restate)
Goal 3 (Reasoning). Encourage students to deepen their own reasoning	
a	Press for reasoning or ask for evidence from the student who is currently speaking "Why do you think that?", "What is your evidence?" (Why)
b	Challenge the student who is currently speaking to elicit more information "Does it always work that way?" (Challenge)
Goal 4 (Thinking with others). Encourage students to work with each other's idea	
a	Prompt students to extend the ideas of other students "Who can add to that?" (Add on)
b	Ask students to explain what someone else said "Who can explain what she means?" (Explain other)
c	Ask students to apply their reasoning to another's reasoning "Do you agree or disagree, and why?" (Agree/disagree)

O'Connor 2015). Following Michaels and O'Connor (2015), we define *productive classroom talk* as a teacher-guided classroom talk that accomplishes four goals (see Table 1): (1) Students share, expand, and clarify their own ideas, (2) students listen to one another, (3) students deepen their own reasoning, and (4) students work with each others' ideas. By *orchestrating (productive) classroom talk*, Michaels and O'Connor (2015) refer to the teacher's initiation of the classroom talk and his/her interaction with the students, so that the talk accomplishes the four goals. The purpose of these goals is to ensure that classroom talk becomes and remains cognitively demanding and therefore productive for student learning (Resnick et al. 2010; Stein et al. 2008). It is precisely the cognitive demand of mathematical tasks that is the content-dependent dimension of instructional quality with crucial impact on student performance (Praetorius et al. 2018).

In a productive classroom talk, the teacher encourages students to participate in mathematical discourse practices. Students benefit from seeing how others approach a task and, therefore, gain insight into others' problem-solving strategies and thought processes that they might not have considered themselves. They learn to make their thought processes explicit, to explain and justify them, and to make connections between concepts, strategies, and representations (for a review, see Resnick et al. 2015). Thus, based on these assumptions, it is expected that the more productive the classroom talk, the more students will provide explanations and justifications for the correct use of concepts, strategies, and representations and link these together. In other words, in productive classroom talk, students can collaboratively construct an appropriate understanding of mathematical concepts and of the relations between concepts (Bauersfeld 1988; Yackel and Cobb 1996). However, it is challenging for teachers to successfully orchestrate this type of classroom talk (Erath et al. 2021, Mehan 1979).

To support teachers in orchestrating productive classroom talks, concrete conversational strategies called *productive talk moves* (Michaels and O'Connor 2015) were developed for teachers. Teachers can use these productive talk moves in classroom talk to encourage students to elaborate, listen, deepen their ideas, and think together with others (see Table 1). By applying productive talk moves appropriately, students' processes of negotiating meaning can be supported. For example, productive talk moves like "Say more" and "Time to think" (see Table 1) are intended to encourage students to describe, explain, and justify their mathematical thinking in detail, and moves such as "Why" and "Challenge" (see Table 1) can specifically initiate the negotiation of meaning and validity in the sense of what can be inferred from what. These are necessary steps towards understanding what the mathematical community endorses and does not endorse as a valid reason for accepting claims.

2.1.2 Use of Productive Talk Moves

Frequencies of productive talk moves used in classroom talk have been documented in intervention studies in which an experimental group is compared with a control group. From these studies, it can be concluded that the majority of teacher turns are not productive talk moves. For example, in an intervention study conducted in the United States on the use of productive talk moves in classroom talk (Grade 5), about

20% of teacher turns before the intervention were productive talk moves (Michaels and O'Connor 2015). After the intervention, about 40% of teacher turns were categorized as productive talk moves. Comparable frequencies have been documented in mathematics classes in Grade 6 and 7 in China (Chen et al. 2020) and in kindergarten science classes in the Netherlands (van der Veen et al. 2017). So far, there are no findings on the use of productive talk moves in elementary schools in Switzerland.

In exploring the use of productive talk moves by teachers, it would be interesting to know whether teachers use each productive talk move comparably often, or whether some teachers use certain moves more often than others. The few findings available indicate that, in the United States, teachers seem to use certain productive talk moves more than others. For example, Michaels and O'Connor (2015) found that elaborating moves (see Table 1) occurred most frequently and listening moves least frequently in Grade 4 and 5. Similarly, the study by van der Veen et al. (2021) presents four teachers who all use very few listening and thinking with others moves prior to attending professional development on productive classroom talk. However, these teachers differ distinctly (from a descriptive perspective) in the number of reasoning moves used, and less distinctly in the number of elaborating moves. Tabach et al. (2020) found a comparable pattern. They demonstrated in a case study of two teachers (Grade 8, Israel) that they differed only in the number of reasoning moves.

From the theoretical perspective, the use of productive talk moves does not necessarily lead to productive classroom talk. The talk sequence is only productive when the productive talk moves result in an interaction that meets the intended demands of the moves. However, various findings show that the number of productive talk moves can serve as a proxy for the productivity of classroom talk: The number of productive talk moves is positively related to the number, length, and content quality of student turns as well as to student performance (Chen et al. 2020; O'Connor et al. 2017; Rüede et al. 2023).

2.2 Classroom Talk in Current Elementary School Mathematics Teaching

Current discussions in German-speaking countries focus on how classroom talks can be beneficially integrated into elementary school mathematics instruction (Häsel-Weide and Nührenbörger 2021). As mentioned above, Häsel-Weide and Nührenbörger (2021) propose segmenting mathematics lessons into launch, explore and reflection phases. The importance of this segmentation is emphasized by teachers as well (Schindler 2017).

During the launch phase, students are introduced to the new problems to be solved using classroom talk. In this phase, classroom talk initiates individual and small group work. It should be clarified jointly which mathematical aspects should be paid attention to in the problems to be solved; highlighting, for example, which (type of) justifications will be important and which forms of representation can be used. During the explore phase, students work on the problems individually or in small groups. In this phase, the teacher can use scaffolding, a form of individual, adaptive support that is both comprehension-oriented and structured (Pfister et al. 2015). Finally, during the reflection phase, different learning experiences can be

summarized, discussed, and elaborated using classroom talk. Collaboratively, criteria and norms that were mostly used tacitly in the explore phase should be made explicit, at least in part. In this phase, individually-constructed meanings should be made accessible to all students.

Classroom talks in elementary school mathematics lessons offer young students their first formal contact with mathematical discourse practice. If classroom talks are integrated into the launch and reflection phases, then they have a high chance of becoming productive for learning. For example, if a cognitively demanding problem is presented in the launch phase (e.g., "Which picture represents 5×3? Why?"; see Fig. 2 on the far left), the class must clarify what is given and what is asked for, what should be attended to in the representations presented, and what will need to be justified and how (Resnick et al. 2010). Through these mathematical discourse practices, multiple student interpretations can be negotiated and the classroom talk thereby becomes productive. Accordingly, in the reflection phase, multiple procedures, interpretations, and justifications that students have explored independently can be presented and productively discussed together.

In classroom talks integrated in the launch and reflection phase, students can learn what mathematical reasons are, how they connect mathematical concepts, and how they assign meaning to them. Especially in the practice of justifying, students are required to provide mathematical reasons (Kilpatrick et al. 2001). Therefore, from a theoretical perspective, the teacher's encouragement of students to generate and elaborate on justifications in classroom talk is of particular value. This process makes students' understanding of mathematical concepts explicit. It is important that students experience such meaning-making processes in the launch and reflection phase, because the kind of meaning-making processes which the teacher manifests in the classroom talk can in turn shape meaning-making processes during the more independent explore phase (Wood 1999). As a consequence, productive classroom talk can affect learning not only in the launch and reflection phases but also in the explore phase.

2.3 Student Justifications

In *student justifications*, students explain why something has to be the way it is by "provid[ing] sufficient reason" (Kilpatrick et al. 2001, p. 130). Student justifications can be understood as a component of student reasoning—in Kilpatrick et al. (2001) it is a component of adaptive student reasoning. Proofs (both formal and informal), for example, are justifications, but not every justification is a proof. Student justifications are often logically incomplete and only suggest the source of the real mathematical reason. For elementary school, this is especially relevant as students take their first steps in justifying mathematical ideas and strategies (Bieda and Staples 2020; Thanheiser and Sugimoto 2022).

We see the prototypical student justification as an answer to a why-question in the form of a because-sentence. In this sense, a student justification can include reasons why (1) a certain mathematical idea, strategy or explanation is correct or not, (2) a produced interpretation (e.g., of a symbolically represented term, of a geometric

representation, etc.) is or is not appropriate, (3) a rule, procedure, or definition is or is not correctly applied.

In elementary school, student justifications are usually not produced spontaneously. Krummheuer (1997, p. 29), for example, observed that students "basically just talk about computations" when solving tasks together; justifications did not occur. That is, students must be encouraged to justify their own ideas and strategies and those of others. For example, a teacher can use productive talk moves (e.g., "Why does this strategy work?", see goal 3 in Table 1) or design situations in which students are productively irritated (Schwarzkopf 2015): unexpected, surprising, or disappointing assertions can evoke spontaneous student justifications. By being asked to justify why their strategies work and why their ideas are correct, students can develop understanding and construct meaning for mathematical concepts and procedures. This was shown, for example, in the case study of Tabach et al. (2020) comparing two teachers (grade 8) who each implemented the same teaching material. From a descriptive perspective, the use of productive talk moves was comparable for both teachers except for the reasoning moves. The second teacher used the reasoning moves (such as "Why"-moves) more often. Consistent with theoretical expectations, the students of this second teacher who utilized reasoning moves increased (descriptively) in their conceptual knowledge more than students of the first teacher; in particular, students of the second teacher produced more justifications.

2.4 Effects of Orchestrating Productive Classroom Talk On Student Justifications

Approaches to the productive orchestration of classroom talk argue that student justifications occur often in productive classroom talk (Alexander 2020; Michaels and O'Connor 2015). This is because in productive classroom talk, students are explicitly asked to justify their own ideas as well as those of others. For example, the productive talk move "Why" (see Table 1, goal 3) can be used to ask students to explain why their ideas are correct. This can directly result in student justifications. Other productive talk moves such as "Time to think", "Explain others", "Agree/disagree" (see Table 1) encourage students to elaborate more on their ideas, comment on others' reasoning, and can likewise prompt the production of student justifications. In a classroom talk, faulty student justifications should be identified and discussed together to see what can be learned from them, and in the case of incomplete student justifications, an attempt is made to collectively fill in the gaps.

Empirical findings confirm these theoretical assumptions (for a review, see Resnick et al. 2015). Quantitative studies often measure the productivity of classroom talks by determining the number of productive talk moves the teachers used. Positive relationships have been found between the number of productive talk moves and, for example, the length and frequency of student turns as well as the number of student reasonings (e.g., Chen et al. 2020; O'Connor et al. 2017). Effects of teachers' orchestration of productive classroom talk specifically on student justifications (and how these justifications influence student performance) have rarely been analyzed (for an exception, see Mok et al. 2022).

Most of the above findings come from studies at the secondary school level. Some studies at the elementary school level do exist (e.g., Häsel-Weide and Nührenbörger 2021; Krummheuer 2007; Steinbring 2000), but the majority focus on the reconstruction of mathematical discourse practices. Studies have not yet been conducted to study quantitative relationships between teachers' orchestration of classroom talk and the quality of students' contributions (such as, e.g., the number of student justifications).

2.5 Effects of Orchestrating Productive Classroom Talk On Student Learning

From a theoretical perspective, relationships are hypothesized between the orchestration of productive classroom talk and student learning. First, it can be argued that when students are asked to justify, they need to make connections between concepts, properties, and examples. This can lead to improved organization of their conceptual knowledge (Chi and Wylie 2014). In productive classroom talk, students are also obliged to make the connections between concepts explicit. Making something explicit is an act of self-explanation (Chi et al. 1994). Second, it can be argued that when students elaborate on their own ideas and are listening to other students and thinking about their responses, they gain insight into solution strategies, interpretations, and reasoning processes that they may not have considered. This allows them to revise and build on their own knowledge (Webb et al. 2014). For example, elaborating on one's own ideas requires students to transform what they know into communication that is coherent and complete. This allows them to recognize misconceptions or incompleteness in their own ideas more than they would when simply verbalizing aloud to oneself (Whitebread et al. 2007). This may encourage students to engage in processes that promote their learning, including seeking new information to correct misconceptions and to fill in gaps in their own understanding.

There are recent studies that have demonstrated the theoretically expected effects of orchestrating productive classroom talk on student performance in mathematics (Chen et al. 2020; O'Connor et al. 2017; Rüede et al. 2023), mostly in secondary school. However, there are also some studies that found no effects on student performance (e.g., in science and mathematics: Pehmer et al. 2015).

At the elementary school level, there are only a few studies. Two examples are the case studies of Hiebert and Wearne (1993: Grade 2) and Murata et al. (2017: Grade 1). Both studies contrasted cases of teachers who orchestrated productive classroom talk with teachers who, for the most part, called for simple facts and procedures in classroom talk and provided justifications, if any, themselves and in a generic way. In both studies, student performance was higher in classrooms with teachers who orchestrated classroom talk productively. However, these case studies are limited to a descriptive perspective. Another example is research on teachers' practices (in Grades 2–4) in encouraging students to engage with each others' ideas (Ing et al. 2015; Webb et al. 2014). These studies analyzed whole-class and small-group discussions. The analyses show that both explaining one's ideas and engaging with other students' ideas at a high level is positively related to student performance. However, it remains open whether whole-class and small-group discussions contribute differently to these effects.

In the studies above, it is not always documented whether the lessons were segmented into launch, explore, and reflection phases. In Hiebert and Wearne (1993), each lesson was segmented into a number of short launch-explore-reflection sequences; Murata et al. (2017) analyzed classroom talks (15–20 min) which were designed specifically to support students' mathematical discourse practices and were offered twice a week; Ing et al. (2015) and Webb et al. (2014) do not specify the segmentation of the lessons analyzed. That is, overall, results on classroom talk seem to come from varied lesson designs. We expect similar positive results for classroom talk in the launch and reflection phases of mathematics lessons in German-speaking elementary schools.

2.6 Research Questions and Hypotheses

Orchestrating productive classroom talk is widely recognized as a vehicle for enabling students to participate in mathematical meaning-making processes: As students elaborate on and deepen their understanding of their own ideas and those of others, multiple meanings of concepts become explicit and are negotiated, supporting meaning construction (Moschkovich 2007, 2015).

Consistent with these assumptions, the positive effects of productive classroom talk on student justifications and student performance have been recently demonstrated (see 2.4 and 2.5). However, the majority of the findings were shown in English-speaking countries and secondary schools. There are no robust findings on the orchestration of productive classroom talk and its relationships with the number of student justifications and with student performance in elementary school level mathematics teaching in German-speaking countries and especially in Switzerland. The present study addresses this research gap.

We used an existing data set (see 3.1) that included mathematics lessons from second grade classrooms in Switzerland. These lessons were segmented into a launch, explore, and reflection phase with classroom talks in the launch and reflection phase. We analyzed the orchestration of the classroom talks in order to investigate the following research questions:

1. *How often do teachers use productive talk moves?*

   Studies of mathematics education have shown that among teachers who have not attended any professional development on productive classroom talk, about 20% of teacher turns can be classified as productive talk moves (see 2.1). These studies refer to different countries and different grade levels (Chen et al. 2020; Michaels and O'Connor 2015; van der Veen et al. 2017). Therefore, we also expected that about 20% of the teacher turns in classroom talk of the launch and explore phase in second grade mathematics classes in Switzerland would be categorized as productive talk moves (hypothesis 1).

2. *What patterns of productive talk moves used can be identified?*

   With regard to the second question, we investigated whether some productive talk moves were used more often by some teachers than others. In line with Tabach

et al. (2020) and van der Veen et al. (2021) we expected that some teachers will use "Why"-moves more often than other teachers (hypothesis 2). However, there are no robust findings on further differential differences in the use of productive talk moves. To complement hypothesis 2, we exploratively analyzed our data for additional patterns of productive talk moves used.

3. *How are the number of productive talk moves related to the number of student justifications?*

Positive effects of the number of productive talk moves on the number of student reasonings have been documented for classroom talk at the secondary school level, in a few cases also in kindergarten (see 2.4). Student justifications are a component of student reasonings, therefore, we expected similar effects for classroom talk in second grade mathematics classes in elementary school in Switzerland: The number of productive talk moves will be positively related to the number of student justifications (hypothesis 3).

4. *How are the number of productive talk moves related to student performance?*

From a theoretical perspective, a positive relationship is expected between the number of productive talk moves in classroom talk and student performance (see 2.5). At the secondary school level, the majority of the empirical findings support this positive relationship (see 2.5). There are considerably fewer findings at the elementary school level, indicating a positive relationship between the number of productive talk moves and student performance. Therefore, we hypothesized that the number of productive talk moves will be positively related to student performance (hypothesis 4).

3 Methods

To investigate research questions 1–4, we used quantitative analyses. Because few quantitative studies have been conducted with the grade level studied and the research questions posed, our analyses also have an explorative character. Therefore, we additionally use three transcripts to present the results of research question 2 (patterns of productive talk moves used).

3.1 Data Collection

We used an existing dataset from the MALKA[2] intervention study (Florin 2021) to improve arithmetic mastery in the first two years of schooling. The present study used data from the second school year, summer 2019 to summer 2020. Overall mathematics performance was assessed in summer 2019 and at the end of the second

---

[2] MALKA is the acronym for "Mathematik lernen und kooperieren von Anfang an" (Learning mathematics and cooperating from the beginning).

school year in summer 2020. During the second school year audio recordings (and student data) were collected.

The aim of the intervention in this second school year was to examine effects of 16 support lessons to foster operation sense. The study included an intervention and a control group. Teachers in the intervention group were provided with teaching material for the 16 support lessons (see 3.3), divided into launch, explore, and reflection phases. At the beginning of the school year 2019/20, they were introduced to the use of the teaching material in a half-day kick-off meeting. Around the middle of the school year, each teacher was contacted by phone to clarify any questions that arose during the implementation of the teaching material. The 16 support lessons were intended to be spread over the school year and to replace 16 regular lessons in mathematics. Thus, the intervention classes did not receive a higher number of mathematics lessons.

The intervention included the implementation of 16 support lessons to foster operation sense. The control group was a business-as-usual control group. To analyze implementation fidelity in the intervention classes, support lesson 11 was recorded[3]. No lesson was recorded in the control group because the control teachers were provided with the teaching material only after the data collection. The teachers were invited to submit an audio recording of the launch and the reflection phase of support lesson 11 (presented in detail in Sect. 3.3), because these phases were designed as classroom talk. The recording can be seen as representative of the teachers' implementation of the support lessons because they were instructed to teach this lesson just as they had taught the other support lessons.

3.2 Participants

Figure 1 outlines the structure of the second year of the longitudinal study, starting in summer 2019 with 68 primary school classes (grade 2); all teachers participated voluntarily. The intervention group consisted of 52 classes; the control group contained 16 classes.

Due to SARS-CoV-2, schools were closed for six weeks in spring 2020. Many teachers could no longer teach regularly. This led to the decision to keep only those teachers in the intervention group who could implement at least the first 11 support lessons. The sample was reduced to 46 classes: 34 (all female[4] teachers) in the intervention group, 12 (all female teachers) in the control group. Only 22 of the 34 teachers in the intervention group had a recording of support lesson 11. These 22 teachers from the intervention group formed the sample for the analysis presented here. The final data set included 22 audio recordings of support lesson 11 and the pretest and posttest scores of a total of 305 students in the 22 classes.

---

[3] Support lesson 11 was chosen because it was identified by three experts in mathematics education as the prototypical example of the 16 support lessons provided. In addition, the time point was taken into account: the teachers should already have gained some experience in implementing the support lessons, but the selected support lesson should also not be at the very end.

[4] In Switzerland, approximately 85% of elementary school teachers are female.
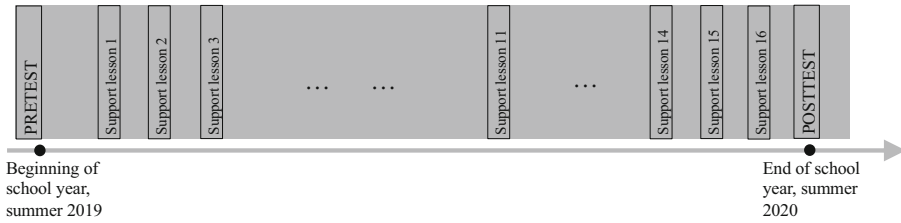
Fig. 1 Timeline of data collection and support lessons in the intervention classes. Support lesson 11 (presented in detail in Sect. 3.3) was recorded

In the MALKA study, control and intervention groups were integrated (for the main results, see Florin 2021). However, the study presented here only uses the data from the intervention group. This is because (1) only support lesson 11 in the intervention group was recorded and (2) the research question investigated here is distinct from the main questions of MALKA.

3.3 Teaching Material for the Support Lessons

All 16 support lessons were designed for elementary school students in the second grade (Florin 2021). Each lesson consisted of one prepared teaching sequence of approximately 30 to 45 min and was designed for additional support in developing operation sense (Royar 2013). The launch and reflection phases of each support lesson included classroom talk, whereas the explore phase was designed as a cooperative setting of pairs of students. For each lesson, the instructional material included specific guidelines (e.g., important questions to ask in classroom talk), notes for implementation, and all representations (in paper and digital) to be shown in the lessons.

We describe support lesson 11 in more detail as an example of the provided content. As stated above, the main points of the learning subject of this support lesson had already been discussed in regular lessons and the support lesson offered additional learning opportunities for a deeper understanding of this subject. In the launch phase, using classroom talk, students are reminded of multiplicative and non-multiplicative structures in iconic representations. The students are asked to evaluate different iconic representations (for examples see Fig. 2, left images) as to whether they can be adequately interpreted as 5×3 or not, and why this is the case. Some of the situations shown are not expected by the students when multiplying 5×3, such
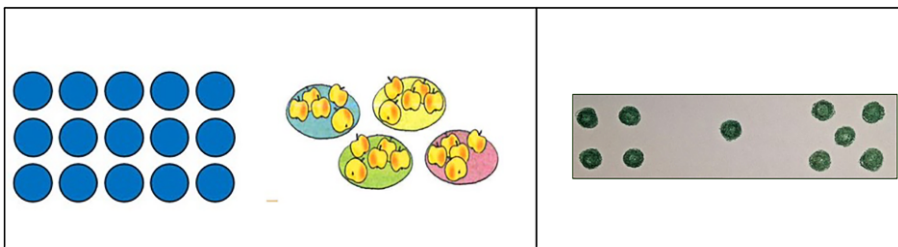
Fig. 2 Examples of representations for use in support lesson 11. The left two representations were examples for use in the launch phase, and the right representation was for the reflection phase (Florin 2021)

as the image of four plates with five apples each. This could "productively irritate" the students (Schwarzkopf 2015, p. 39) and therefore motivate them to justify why four plates with five apples each can or cannot be adequately interpreted as 5×3. To support the orchestration of classroom talk, examples for helpful teacher turns are provided in the teaching material, e.g., "Can you explain why the picture fits or does not fit?".

In the explore phase of support lesson 11, students are asked to elaborate on further iconic representations in pairs. The support lesson concludes with a reflection phase to discuss in depth important findings that the students had gained in the explore phase. To initiate this classroom talk the teacher should use a turn such as "Hanna (a child from another school) thinks this picture fits 4×5" (see Fig. 2, right side). This prompt asks students to verbalize criteria they can use to evaluate the appropriateness of iconic representations of 4×5. To do this, students must recall their learning experiences from the explore phase and make explicit the criteria tacitly used in that phase.

3.4 Measures

We used deductively derived categories to code teacher and student turns, and we measured student performance with a standardized written test.

3.4.1 Coding of Teacher Turns: Productive Talk Moves

To determine the number of productive talk moves, we analyzed the audio recordings of support lesson 11. In the recordings, we identified the launch and reflection

Table 2 Codes for teacher turns. On the left the labels of the codes are given, in the middle their description and on the right examples from the transcripts

	PRODUCTIVE[a]	
PRODUCTIVE Say more	"Say more" moves	"Can you tell me more about it?"
⋮	⋮	
PRODUCTIVE Why	"Why" moves	"Can you explain why?" "Can you show me?"
⋮	⋮	
PRODUCTIVE Agree/disagree	"Agree/disagree" moves	"Would you do anything differently than Emma? Why?"
	NON_PRODUCTIVE	
ADMIN	Teacher turns addressing only lesson organization	"When you both have completed your sheet, you will sit down together" "Yes, you may go to the toilet"
CONTINUER	Teacher turns aiming solely at maintaining and directing the discourse (Gardner, 2001)	"Okay", "Anna", "Mmh, Chantal?"
OTHER	All other teacher turns	"Here you can see four times five points" "Right, three times five fits here"

[a]We used all ten subcodes of Table 1

phase and transcribed them according to a transcription manual. Then teacher and student turns were coded by the same raters, both with expertise in mathematics education. The codes were chosen so that we could subsequently categorize each turn as either productive or non-productive. For identifying productive talk moves we used the (sub)goals of productive classroom talk (Table 1; O'Connor et al. 2017 and Chen et al. 2020). If a teacher turn addressed one of these (sub)goals and an arithmetic subject, we coded it as PRODUCTIVE, see Table 2. We coded a teacher turn only then as PRODUCTIVE when the intention of the turn was made explicit. For example, a teacher turn like "You think so?" (see Table 8, turn 3) may result in the student saying more. However, "You think so?" does not explicitly ask the student to say more, so we did not code this turn as "PRODUCTIVE Say more".

To distinguish between the non-productive teacher turns we used three additional subcodes: We differentiated between teacher turns addressing lesson organization (ADMIN), helping to support discourse (CONTINUER), and all others (OTHER; see Table 2).

Based on these codes, all teacher turns (total 895) were coded. We double-coded 27% of the data (six of 22 transcripts). The interrater reliability was almost perfect ($\kappa = 0.81$; Landis and Koch 1977).

3.4.2 Coding of Student Turns: Student Justifications

We assigned exactly one code to each student turn: either a student turn was categorized as justification (code JUSTIFICATION) or it was not (code NO_JUSTIFICATION). We based our codes on the teacher turns that were previously coded, especially on the "Why"-moves. To achieve high interrater reliability, we distinguished several subcodes (see Table 3). A student answer to a "Why"-

Table 3 Codes for student turns. On the left the labels of the codes are given, in the middle their description and on the right examples from the transcripts

	JUSTIFICATION	
WHY_BECAUSE	Student answers to a why question including syntactical features ("because", "for", "therefore", etc.) and semantically including reasons	"This doesn't fit, because it's four time five" as an answer to "Why?"
WHY_WITHOUT-BECAUSE	Student answers to a why question without syntactical features but semantically including reasons	"There are too many, there are about twenty" to the teacher question "Why not?"
WHITHOUT-WHY_BECAUSE	Student turns including syntactical features indicating a reason without having been asked for	"It does not fit because there are four plates, then there are twenty"
EVIDENCE	Student turns that demonstrate the chosen interpretation, if it was asked for	"Here's the first row, second, third, fourth, and fifth" to the teacher question "Can you show it?"
	NO_JUSTIFICATION	
ADMIN	Student turns addressing only lesson organization	"Can we take a pencil, too?" "Who am I with?"
OTHER	All other student turns	"Five" "There are three times five"

move that addressed an arithmetic subject was coded as WHY_BECAUSE if it included reasons and syntactic features, as WHY_WITHOUT_BECAUSE if it included reasons but no syntactic features, and as EVIDENCE if it demonstrated the chosen interpretation of a representation. A student turn addressing an arithmetic subject that was not an answer to a "Why"-move was coded as WHITOUT-WHY_BECAUSE if it included reasons and syntactic features. Again, we used the codes ADMIN and OTHER for student turns that did not include a justification (see Table 3).

Using these codes, we coded all 22 transcripts (898 student turns). We double-coded 27% of the data (six transcripts). The interrater reliability was substantial ($\kappa = 0.79$; Landis and Koch 1977).

3.4.3 Student Performance

Student performance was assessed using the BASIS-MATH. This standardized test measures the overall arithmetic knowledge of students and is adapted to the Swiss elementary school. It also makes it possible to differentiate between children with lower mathematical abilities, which was a specific focus in the MALKA intervention study (Florin 2021).

Before the beginning of grade 2 (pretest) the BASIS-MATH 1+ was used (28 items, maximum test score 28, Schnepel et al. in preparation), at the end of grade 2 (posttest) the BASIS-MATH 2+ (30 items, maximum test score 30, Moser Opitz et al. 2020).

The BASIS-MATH 1+ tests students' understanding and skills with respect to numbers, addition, subtraction, and story problems and the BASIS-MATH 2+ with respect to place value, addition, subtraction, multiplication, and story problems. Reliability was high for both tests, 0.89 (pretest) and 0.89 (posttest).

3.5 Data Analyses

Due to the inclusion criteria for the final sample, there was no missing data for audio recordings or BASIS-MATH test scores.

Descriptive statistics were used to test hypothesis 1 (research question 1, number of productive talk moves). To answer research question 2, a hierarchical cluster analysis using Ward's method with Euclidean distance as a measure of similarity was conducted (Borgen and Barnett 1987). Hierarchical algorithms such as Ward's are useful for exploratory analysis when the likely number of clusters is not fixed in advance (Antonenko et al. 2012). The clustering variables were the number of elaborating moves, listening moves, reasoning moves, and thinking with others moves. Finally, to present the cluster analysis, we selected a transcript for each pattern that characterized how classroom talk of that pattern looks. To test hypothesis 3, a linear regression of the number of student justifications on the number of productive talk moves was calculated (research question 3, relationship between number of productive talk moves and number of student justifications). For the research questions 1, 2, and 3 we used SPSS 27.

To answer research question 4 (relationship between number of productive talk moves and student performance), we used hierarchical multilevel models, because students were nested within classes. Our multilevel model included student-related variables ($Posttest_{ij}$, $Pretest_{ij}$) at the student level and class- or teacher-related variables ($Pretest\_mean_j$, $Productive_j$) at the class level. In our model $Pretest_{ij}$ denotes the pretest score of student i in class j and $Posttest_{ij}$ denotes the posttest score of student i in class j. $Pretest\_mean_j$ denotes the mean of the pretest scores and $Productive_j$ the number of productive talk moves in class j. In the following, Eq. 1, 2, and 3 describe our model, and the residuals are denoted with r_{ij}, u_{0j}, and u_{1j}.

Student-Level:

$$Posttest_{ij} = \beta_{0j} + \beta_{1j} \times Pretest_{ij} + r_{ij} \tag{1}$$

Class-Level:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} \times Pretest\_mean_j + \gamma_{02} \times Productive_j + u_{0j} \tag{2}$$

$$\beta_{1j} = \gamma_{10} + u_{1j} \tag{3}$$

By using group-mean centering for the students' pretest performance scores and grand-mean centering for the pretest performance scores of the classes, we were able to separate the influence of the pretest performance scores into a part between classes (γ_{01}) and a part within the students of a class (γ_{10}) (Enders and Tofighi 2007). For our multilevel models we used the two-level function and full maximum likelihood estimation in MPlus 8.

4 Results

A total of 895 teacher turns and 898 student turns were analyzed, that is, in total 1793 turns. Twenty-one teacher turns and 23 student turns were coded as ADMIN. We excluded these turns from the analyses because we were interested in turns addressing arithmetic subjects. We therefore analyzed 1749 turns: 874 teacher turns and 875 student turns.

Table 4 shows the descriptive statistics of the preliminary analyses of the audio recordings. On average, slightly more than 9 min of classroom talk were integrated into the support lesson 11. From a descriptive perspective, students talked as often as teachers, but teachers talked longer. Overall, variances were high, suggesting differences between teachers in how classroom talk was integrated into the support lesson 11 (see Table 5).

Table 4 Means and standard deviations (in parentheses) of length of classroom talk and of length and number of turns (student, teacher, total)

Length of classroom talk	Length of teacher turns	Length of student turns	Number of teacher turns	Number of student turns	Number of total turns
9.12 min (4.18 min)	8.50 s (2.13 s)	5.81 s (3.13 s)	39.73 (17.21)	39.77 (18.77)	79.50 (35.78)

Table 5 Absolute numbers of productive talk moves per class (elaborating, listening, reasoning, and thinking with others moves) and of teacher turns. For the numbers of total productive talk moves, the relative numbers (per number of teacher turns in the class) are given in parentheses

| Teacher	Elaborating moves	Listening moves	Reasoning moves	Thinking with others moves	Total productive talk moves (%)	Teacher turns
A	9	0	2	8	19 (30.3)	62
B	4	0	7	0	11 (28.2)	39
C	1	0	0	0	1 (4.5)	22
D	6	0	9	1	16 (27.1)	59
E	2	0	10	0	12 (17.6)	68
F	2	0	4	0	6 (12.5)	48
G	1	0	10	0	11 (18.6)	59
H	2	0	11	2	15 (35.7)	42
I	1	0	1	0	2 (25.0)	8
J	2	0	1	0	3 (12.5)	24
K	6	0	7	1	14 (43.8)	32
L	2	0	1	0	3 (10.3)	29
M	2	0	3	1	6 (12.5)	48
N	2	0	1	1	4 (12.9)	31
O	5	0	4	0	9 (18.0)	50
P	2	0	7	1	10 (33.3)	30
Q	0	0	4	0	4 (18.2)	22
R	1	0	1	0	2 (8.0)	25
S	4	0	9	1	14 (21.9)	64
T	1	0	5	0	6 (10.3)	58
U	2	0	4	0	6 (15.8)	38
V	1	0	2	0	3 (18.8)	16
Total	58	0	103	16	177 (20.3)	874
% of productive talk moves	32.8%	0.0%	58.2%	9.0%	–	–

4.1 Productive Talk Moves

Our first research aim was to investigate which productive talk moves teachers used in orchestrating classroom talk and how often they used them. Table 5 shows the descriptive statistics of the numbers of applied productive talk moves.

In line with Hypothesis 1, on average about 20% of the teacher turns were productive talk moves. From a descriptive perspective, teachers used reasoning moves the most, followed by elaborating moves. Thinking with others moves were rather rare and listening moves did not occur. Some teachers used very few productive talk moves. For example, teacher C used only one productive talk move (in a total of 22 teacher turns).

4.2 Patterns of Productive Talk Moves Used

The second aim of our study was the identification of patterns of teachers' productive talk moves. For the cluster analysis we used the number of elaborating moves, reasoning moves, and thinking with others moves. Listening moves were omitted because they were not used by any teacher (see Table 5). The dendrogram in Fig. 3 visualizes the result of the cluster analysis. It reads from left to right and describes the process of clustering in this direction. Initially, each teacher is in his/her own cluster, as indicated by each teacher having his/her "own" short horizontal line, in total 22 short horizontal lines. These clusters are gradually merged into larger clusters from left to right. Vertical lines illustrate that two clusters are merged.

The horizontal axis describes heterogeneity, normalized to the range from 0 to 25. Heterogeneity increases when two clusters are merged. The more it increases, the more inhomogeneous the newly joined clusters are. The number of clusters is determined by the largest increase in heterogeneity in the dendrogram (Borgen and Barnett 1987). In Fig. 3, this is between the three-cluster solution (heterogeneity 5) and the two-cluster solution (heterogeneity 14). That is, the reduction from three clusters to two clusters makes the two resulting clusters highly heterogeneous suggesting a three-cluster solution. Table 6 shows the three clusters (three patterns of
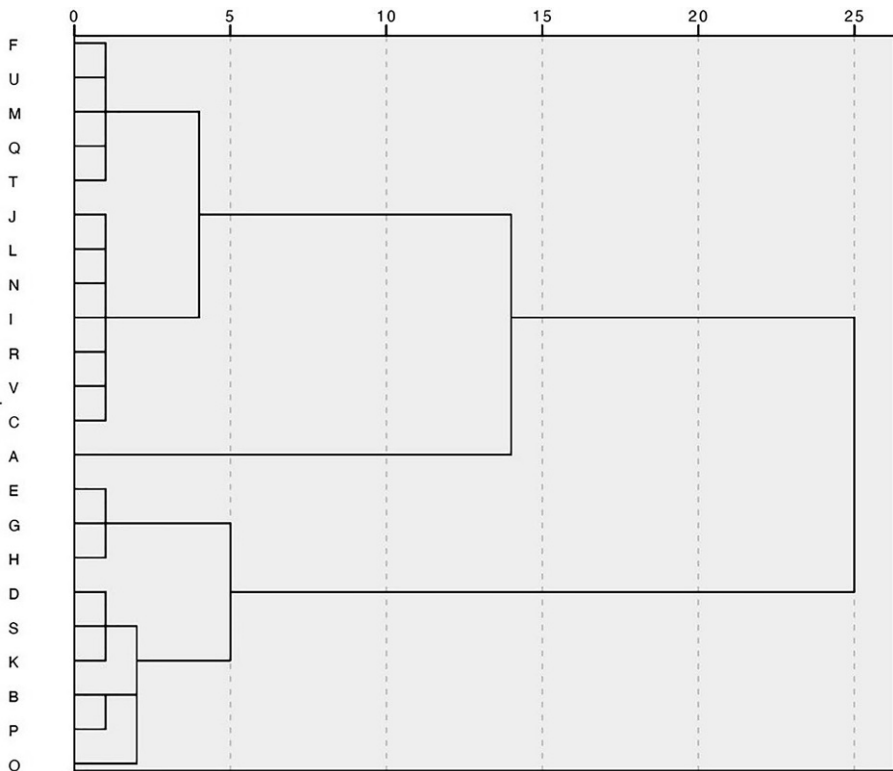
Fig. 3 Dendrogram of the cluster analysis

Table 6 Features of three patterns of productive talk moves used

	Pattern 1 (Cluster 1)	Pattern 2 (Cluster 2)	Pattern 3 (Cluster 3)
Feature of productive talk moves used	Many reasoning moves	Many elaborating and thinking with others moves	Few productive talk moves
Feature of classroom talk	Ideas have to be justified	Collaborative justifying	Few justifications
Assigned teachers (number)	B, D, E, G, H, K, O, P, S (9)	A (1)	C, F, I, J, L, M, N, Q, R, T, U, V (12)
Mean of elaborating/ reasoning/thinking with others moves	3.6/8.2/0.7	9/2/8	1.4/2.3/0.2

productive talk moves used). In line with hypothesis 2, the teachers differ in their patterns of reasoning moves used. In addition to hypothesis 2, our sample contains two patterns with few reasoning moves (and not one only).

4.2.1 Pattern 1: Many Reasoning Moves

We present a transcript from the launch phase for each pattern to characterize how teachers using that pattern orchestrated classroom talk. We start with pattern 1 (see Table 6). The teachers with pattern 1 used many reasoning moves, comparatively few elaborating moves, and almost no thinking with others moves. A typical example is teacher H, who used the most reasoning moves of all teachers, 11 in total (see Table 5).

Table 7 shows a transcript fragment from teacher H, the beginning of the launch phase of support lesson 11. Two productive talk moves are reasoning moves (two PRODUCTIVE Why), one is an elaborating move (PRODUCTIVE Wait time).

Table 7 Transcript fragment of T_H's orchestration of classroom talk

Turn number	Speaker	Turn	Code
1	T_H:	You can see a multiplication and I have four pictures on the board. Your task is to think about which one of these pictures really fits the problem. *(Waiting 11 s)* What do you think of this picture? Does it fit the multiplication? *(Waiting 5 s)* S1?	**PRODUCTIVE Wait time**
2	S1:	Yes	OTHER
3	T_H:	Why do you think this rectangular arrangement of tiles fits the multiplication? Can you explain why?	**PRODUCTIVE Why**
4	S1:	Because there are five times um these groups of three	WHY_BECAUSE
5	T_H:	Can you show me?	**PRODUCTIVE Why**
6	S1:	There is a group of three, there and there and there and there	EVIDENCE
7	T_H:	Ah, very well explained. So, this arrangement of tiles fits the problem five times three. What do you think of the picture with the apples? Does it fit the multiplication? S2	OTHER

The productive talk move codes are printed in bold

T_H teacher H, S student(s)

Teacher H starts with the question (turn 1) of whether the representations fit the multiplication 5×3, especially the rectangular arrangement of the three by five tiles (see Fig. 2 on the far left). After a student affirms, she asks for reasons (turn 3) and requested evidence (turn 5), which led directly to student justifications. Overall, teacher H's example shows that through the use of reasoning moves, students must and do justify. Structures are made explicit, first verbally and then through gestures. The focus on reasoning moves, however, leads students to refer little to what other students have said.

4.2.2 Pattern 2: Many Elaborating and Thinking with Others Moves

This pattern includes only teacher A, an outlier in the analyzed sample (see Table 5). Teacher A used many elaborating and thinking with others moves compared to reasoning moves.

Table 8 Transcript fragment of T_A's orchestration of classroom talk

Turn number	Speaker	Turn	Code
1	T_A:	As you can see there is a multiplication. I brought along different pictures to it I want you to look at. You tell me: does it fit and decide, or um explain why it fits. Does this picture fit or does it not fit the problem 5 times 3? You can see the first picture here. *(Waiting 11 s)* Yes S1?	**PRODUCTIVE Wait time**
2	S1:	It fits	OTHER
3	T_A:	You think so?	OTHER
4	S1:	Because it has three rows and um five columns. Like five, um from right to left and three from top to bottom	WHITHOUT-WHY_BECAUSE
5	T_A:	You try to tell us how it fits. Would you like to draw it in the picture?	**PRODUCTIVE Why**
6	S1:	Right here is five times three *(drawing 13 s)*	EVIDENCE
7	T_A:	What did you do?	OTHER
8	S1:	I made five times three	OTHER
9	T_A:	Uh-huh, can you tell me more about it?	**PRODUCTIVE Say more**
10	S1:	Yes, yes you can make another	OTHER
11	T_A:	Yes?	OTHER
12	S1:	There three times five	OTHER
13	T_A:	Good, thank you very much	OTHER
14	S1:	Can I draw it in too?	OTHER
15	T_A:	Can anyone explain what S1 did here?	**PRODUCTIVE Explain other**
16	S2:	He did um, draw in three times five, no five times three	OTHER
17	S3:	Three times five is the reverse, it has the same outcome, and one draws it in differently. It's just the other way around, like horizontally. But it's not just three dots but five	OTHER

The productive talk move codes are printed in bold

T_A teacher A, S student(s)

On the first turn, teacher A waited 11 s (see Table 8). This long waiting time may have prepared S1 to provide the justification immediately in turn 4. This justification made clear the structure to be established in the rectangular arrangement of the three by five tiles (see Fig. 2, picture on the far left). In order to make the structure apparent, teacher A asked student S1 to mark it in the picture (turn 5, asking for evidence, coded as PRODUCTIVE Why, see 3.4.1). Overall, the structure was described verbally (turn 4) as "three rows, five columns" and as "the five from right to left and three from top to bottom" and made visible through markings (turn 6). This seemed so important to the teacher that she asked the student to say more about this (turn 9). In response, student S1 reframed the arrangement of the tiles and interpreted it as "three times five" (turn 12). At this moment, teacher A did not seem to be interested in the reinterpretation, but rather wanted to elaborate on the first structure "five times three." Presumably, she therefore asked all students to explain what S1 had said (turn 15). The other students, however, explained the reinterpretation that S1 had produced (turn 16, 17).

This example addresses interpretations of the 5 on 3 rectangle arrangement of blue tiles as 5×3 (Fig. 2). The teacher's orchestration results in multiple student justifications for why the rectangle arrangement can be interpreted as 5×3. The student justifications show the class that interpretations are to be described, explained, clarified, and justified, manifesting the obligation to give reasons in mathematics.

4.2.3 Pattern 3: Few Productive Talk Moves

Pattern 3 includes teachers who used few productive talk moves overall (see Table 6). A typical example is teacher L. Her transcript fragment demonstrates that students did not give reasons for their answers, nor were they asked to do so (see Table 9). The only subject addressed in this fragment is which multiplication matches which picture. As a result, the interpretations of the pictures were not made explicit in classroom talk. How to read the representations, how to interpret them, and what to pay attention to remained unspoken.

This example shows that student justifications can be absent if the teacher does not productively orchestrate classroom talk.

4.3 Productive Talk Moves and Student Justifications

The third aim of our study was to examine whether the number of teacher's productive talk moves (elaborating, listening, reasoning, thinking with others) was positively related to the number of student justifications. First, descriptive statistics show that means and standard deviations were comparable (student justifications: $M = 8.45$; $SD = 4.67$; productive talk moves: $M = 8.05$, $SD = 5.30$).

A linear regression showed that the number of productive talk moves and the number of student justifications are positively related ($B = 0.59$, $SE = 0.146$, $p < 0.001$) with $r = 0.65$, which is a large effect (Cohen 1992). According to the linear model, one additional productive talk move is connected with an average increase of 0.59 in the number of student justifications. This finding confirms hypothesis 3.

Table 9 Transcript fragment of T_L's orchestration of classroom talk

Turn number	Speaker	Turn	Code
1	T_L	There are different pictures. You will have to decide if it fits the multiplication, you can read here. Does or does it not fit? Yes, S1?	OTHER
2	S1	Four times five	OTHER
3	T_L	Which one is four times five?	OTHER
4	S1	The apples on the plates	OTHER
5	T_L	So, does it fit the problem, or does it not fit?	OTHER
6	S1	Uh-uh *(negative)*	OTHER
7	T_L	*(Waiting 5s)* The apples do not fit. S2?	**PRODUCTIVE Wait time**
8	S2	Well, um there, four times three does not fit either	OTHER
9	T_L	Which one is four times three?	OTHER
10	S2	The one with the water bottles	OTHER
11	T_L	S3?	CONTINUER
12	S3	The tiles, three times five	OTHER
13	T_L	Does it fit?	OTHER
14	S4	Yes	OTHER

The productive talk move codes are printed in bold

T_L Teacher L, S Student(s)

4.4 Productive Talk Moves and Student Performance

The fourth aim of our study was to investigate the relationship between the number of productive talk moves and student performance. Student performance was measured as pretest scores ($M = 20.45$, $SD = 2.80$, maximum test score 28) and posttest scores ($M = 21.32$, $SD = 2.89$, maximum test score 30) of overall arithmetical knowledge.

The Intra Class Correlation (ICC) of student performance at posttest was moderate (0.11). This indicated considerable variances between classrooms, so we ran a step-wise multilevel regression model-building analysis to assess the potential predictors of student performance in the posttest. First, we included individual student performance in the pretest, then pretest class mean, and finally the number of productive talk moves. Model results appear in Table 10. All variables were significant predictors. The analysis confirmed a positive relationship between the

Table 10 Estimated unstandardized regression coefficients for student performance in posttest, standard errors in parentheses

	Model 0	Model 1	Model 2	Model 3
γ_{00} (intercept)	21.33 (0.57) ***	21.30 (0.60) ***	21.31 (0.41) ***	21.31 (0.35) ***
γ_{10} ($Pretest_{ij}$)	–	0.72 (0.05) ***	0.75 (0.05) ***	0.75 (0.05) ***
γ_{01} ($Pretest\_mean_j$)	–	–	0.70 (0.17) ***	1.01 (0.16) ***
γ_{02} ($Productive_j$)	–	–	–	0.24 (0.07) **

* $p < 0.05$

** $p < 0.01$

*** $p < 0.001$

number of productive talk moves and student posttest performance, controlling for student pretest performance at the individual and classroom level. In line with our hypothesis 4, one additional productive talk move was associated with an increase in posttest performance (0.24 points on average). As an additional result, the models in Table 10 also demonstrate strongly positive relationships between student pretest and posttest performance at the individual and classroom level.

We additionally computed a standardized regression coefficient. To do this we multiplied 0.24 by the pooled standard deviation of the number of productive talk moves and divided it by the standard deviation of the posttest scores. The standardized regression coefficients obtained can be transformed into a correlation coefficient (Peterson and Brown 2005), resulting in $r = 0.48$, a medium effect size (Cohen 1992).[5]

5 Discussion

This study investigated the orchestration of classroom talk in second grade mathematics classes in elementary schools in Switzerland. We analyzed classroom talk that came from the launch and reflection phase of a support lesson (support lesson 11). Four hypotheses were tested.

5.1 Number of Productive Talk Moves

Our descriptive data demonstrates that in line with hypothesis 1, approximately 20% of all teacher turns in the Swiss elementary mathematics classroom are productive talk moves. This frequency is comparable to numbers documented for higher grades and in other countries (Chen et al. 2020; Michaels and O'Connor 2015). The variance of productive talk moves in our sample was high: the proportion of productive talk moves used by teachers varied from 4.5 to 43.8%. This is impressive, since all teachers had the same prepared teaching material. They seem to have implemented the teaching material very differently.

In our sample, elaborating and reasoning moves were used in the majority of cases. It is striking that no listening moves and only a few thinking with others moves could be identified in the transcripts. This is in line with findings in Michaels and O'Connor (2015). What this suggests is that in some classroom talks in elementary school, students are rarely asked to collaboratively negotiate and construct mathematical meaning. These students may therefore be unlikely to compare and contrast the ideas of others with their own in classroom talk. Perhaps this lack of listening and thinking with others moves is a consequence of the scarce professional development programs offered in German-speaking countries on how to orchestrate productive classroom talk in elementary school mathematics teaching.

---

[5] Moreover, further analyses showed that the effects in Model 3 are comparable to a model additionally controlling for the length of classroom talk (in seconds).

5.2 Patterns of Productive Talk Moves Used

In our sample, we were able to identify three distinct patterns of using productive talk moves. Some teachers used many reasoning moves (pattern 1), and others used few (pattern 2 and 3) in line with hypothesis 2. In addition to hypothesis 2, we found two different patterns (2 and 3) among those who used few reasoning moves: One teacher used mostly elaborating and thinking with others moves (pattern 2), and the rest used hardly any productive talk moves (pattern 3). Most salient is the pattern 2 of the teacher who was an outlier in our sample. This teacher's way of orchestrating classroom talks could have encouraged students to elaborate on each others' ideas and thus to engage with multiple meanings of mathematical concepts. Notably, this teacher's students showed the highest average learning gains in our sample. This confirms prior findings that suggest encouraging students to participate in collaborative negotiation of meaning can positively affect student performance (Ing et al. 2015; Webb et al. 2014).

Together with the substantial differences in the number of productive talk moves (see 5.1), the three patterns of productive talk moves used show that there was diversity in the orchestration of classroom talk in the analyzed sample. As these classroom talks took place in the launch and reflection phase of support lesson 11, the differences we found suggest that teachers varied greatly in how they integrated the learning opportunities of the explore phase into the lesson. Whether the patterns of productive talk moves used can be reconstructed in other samples and in other subject matter domains will have to be shown in further studies.

5.3 Productive Talk Moves and Student Justifications

The third hypothesis was also confirmed: A positive relationship was found between the number of productive talk moves and the number of student justifications. This confirms findings from higher school levels (Ing et al. 2015; Mok et al. 2022). The result is of value in two ways: first, students at the elementary school level can produce justifications; second, the number of these student justifications is positively related to the productivity of orchestrating classroom talk.

The analyzed student justifications were generated based on interpretations of iconic representations of the multiplication 5×3. This is a widely known task format and can be found in current teaching materials such as the "Schweizer Zahlenbuch" (Kocher et al. 2021). Our result might suggest that for asking for justifications can be used within the everyday tasks already implemented in teaching materials. However, based on our data we cannot determine how often and which task from the mentioned school books were used by teacher. Nevertheless, it is crucial that the tasks are used productively in the classroom. For example, in a launch phase, the teacher can use classroom talk to clarify which justifications and which representations should be focused on when exploring such problems (Häsel-Weide and Nührenbörger 2021; Stein et al. 2008). In a reflection phase, this would give the class the opportunity to make explicit and evaluate criteria and norms that were used tacitly in the exploration phase and that are meaningful for the solution of the task.

5.4 Productive Talk Moves and Student Performance

Our results confirm hypothesis 4. The number of productive talk moves and the student performance were positively related in our sample. This supports findings from higher school levels (Chen et al. 2020; O'Connor et al. 2017; Rüede et al. 2023) suggesting that productive orchestration of classroom talk is related to learning success. This result shows that important findings on the orchestration of productive classroom talk, which are known from studies conducted in secondary schools, seem to be transferable to elementary schools.

The data we relied upon to investigate the relationship between the number of productive talk moves and student performance was rather modest. Further studies will have to investigate this relationship more comprehensively and in more depth in order to generate a robust set of findings.

5.5 Limitations

In the following, we discuss some limitations of our study.

5.5.1 Determination of the Number of Productive Talk Moves

For the analysis of classroom talk we recorded only one lesson (support lesson 11, see 3.2 and 3.3) per teacher, with classroom talks averaging 9.1 min in length. Thus, 9.1 min represents the average total time a teacher orchestrated classroom talk in her lesson addressing an arithmetical learning subject. The first limitation is the short time of 9.1 min and the second limitation is the recording of only one support lesson.

In a short sequence, there may be more or less productive talk moves and student justifications than what would typically occur in classroom talks of this length orchestrated by the teacher. However, studies conducted in elementary schools show that features of classroom talk can be measured even in short sequences (e.g., Decristan et al. 2020, 8.5 min on average, counting student hand-raising; Fishman et al. 2017, 5 to 15 min, determining discourse quality).

In addition, one recorded lesson may not be representative. How many lessons are needed to accurately observe instructional features? Studies differ in how many lessons per teacher they recorded. Some assessed more than one lesson: For example, SINUS for primary schools recorded one to three lessons (Dalehefte and Rieck 2014), and the Pythagoras video study recorded three lessons (Drollinger-Vetter 2011). The well-known TIMSS video study (1999), in contrast, recorded one lesson to measure dimensions of instructional quality (Hiebert et al. 2003). Recent results suggest that content-independent dimensions of instructional quality (e.g., classroom management) can be observed accurately based on one recording (Praetorius et al. 2014). Content-dependent dimensions of instructional quality (e.g., cognitive demand), in contrast, usually cannot be accurately observed based on only one lesson. However, if the content and situational features of the observed lessons are comparable across teachers, then even one lesson can provide an accurate measurement (Praetorius et al. 2014).

The number of productive talk moves can be regarded as a combination of content-independent and -dependent aspects. It appears that one observed lesson may be sufficient to measure differences between teachers in the number of productive talk moves used. A prerequisite for this is that the content and situational features between teachers' lessons are as comparable as possible (Praetorius et al. 2014). The following reasons suggest that this prerequisite may be satisfied for our sample: (1) Each teacher used the same lesson plan for the recorded support lesson, (2) the implementation fidelity of this support lesson was high for each teacher (Streit & Rüede 2023), (3) teachers had already implemented ten support lessons prior to the audio recording, (4) we instructed teachers to teach support lesson 11 in the same way as the first ten support lessons.

5.5.2 Relating the Number of Productive Talk Moves and Student Performance

There are some limitations that affect the results on the relationship between the number of productive talk moves and student performance. (1) We examined the relationship between features observed in 9.1 min of classroom talk in one lesson and the gain in student performance over a full year. We have already discussed in Sect. 5.5.1 above that our results may be limited because we determined the number of productive talk moves using only one lesson. (2) We included only pretest performance (class mean) in the regression models as a control variable at the class level. However, other factors at the class level such as error culture, cognitive demand of written tasks, attitude of the teacher toward the learners, individual support of the teacher, classroom management etc. might also have an impact. Unfortunately, in the MALKA intervention study, from which the analyzed dataset comes, no further variables were collected at the class level (Florin 2021).

Nevertheless, the significant advantage of our data set is that several variables at the class level that vary in other data sets are held constant: (1) As stated above, each teacher implemented the same lesson plan for the recorded lesson with high fidelity (Streit & Rüede 2023), (2) each teacher was female, was part of the intervention group, and implemented at least 11 of 16 support lessons, (3) the teaching materials used in the classes over the whole year were comparable (Florin 2021). Thus, differences in classroom talk are mainly due to differences in its orchestration—and not due to differences in the design of the recorded lesson. Although these points increase the validity of our results, generalizations about the results should be made with caution.

6 Conclusion

The present study shows that in mathematics lessons of second grade classrooms in Switzerland, teachers use productive talk moves in about 20% of their turns when orchestrating classroom talk. However, the differences between teachers are large. Some teachers used mainly reasoning moves, one teacher used few reasoning moves and many elaborating and thinking with others moves, and other teachers hardly used productive talk moves at all.

In addition, for the analyzed sample we show two main findings: First, the number of productive talk moves is positively related to the number of student justifications. Second, student performance in the posttest is positively related to student performance in the pretest and to the number of productive talk moves. To the best of our knowledge, such relationships have only been shown for secondary education and mostly for English-speaking countries. Thus, the present study demonstrates that these positive findings on productive classroom talk may be transferable to the elementary school level and to the German-speaking context, in particular to mathematics teaching in Switzerland.

References

Alexander, R. (2020). *A dialogic teaching companion*. Routledge.

Antonenko, P. D., Toy, S., & Niederhauser, D. S. (2012). Using cluster analysis for data mining in educational technology research. *Educational Technology Research and Development*, *60*(3), 383–398. https://doi.org/10.1007/s11423-012-9235-8.

Barwell, R. (2020). Language background in mathematics education. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 441–446). Springer.

Bauersfeld, H. (1988). Interaction, construction, and knowledge: alternative perspectives for mathematics education. In T. Cooney & D. Grouws (Eds.), *Effective mathematics teaching* (pp. 27–46). Erlbaum.

Bieda, K. N., & Staples, M. (2020). Justification as an equity practice. *Mathematics Teacher: Learning and Teaching PK-12*, *113*(2), 102–108. https://doi.org/10.5951/MTLT.2019.0148.

Borgen, F. H., & Barnett, D. C. (1987). Applying cluster analysis in counseling psychology research. *Journal of Counseling Psychology*, *34*(4), 456–468. https://doi.org/10.1037/0022-0167.34.4.456.

Chen, G., Chan, C. K. K., Chan, K. K. H., Clarke, S. N., & Resnick, L. B. (2020). Efficacy of video-based teacher professional development for increasing classroom discourse and student learning. *Journal of the Learning Sciences*, *29*(4–5), 642–680. https://doi.org/10.1080/10508406.2020.1783269.

Chi, M. T., & Wylie, R. (2014). The ICAP framework: linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823.

Chi, M. T. H., DeLeeuw, N., Chiu, M., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*, 439–477. https://doi.org/10.1016/0364-0213(94)90016-7.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155.

Dalehefte, I. M., & Rieck, K. (2014). How do German primary school teachers prepare students for science standards? Findings from a video study of the professional development program SINUS for primary Schools. *Form@re*, *14*(2), 7–20. https://doi.org/10.13128/formare-15123.

Decristan, J., Fauth, B., Heide, E. L., Locher, F. M., Troll, B., Kurucz, C., & Kunter, M. (2020). Individuelle Beteiligung am Unterrichtsgespräch in Grundschulklassen: Wer ist (nicht) beteiligt und welche Konsequenzen hat das für den Lernerfolg? *Zeitschrift für Pädagogische Psychologie*, *34*, 171–186. https://doi.org/10.1024/1010-0652/a000251.

Deutschschweizer Erziehungsdirektorenkonferenz (2016). *Lehrplan 21 – Mathematik*. D-EDK.

Drollinger-Vetter, B. (2011). *Verstehenselemente und strukturelle Klarheit*. Waxmann.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological Methods*, *12*(2), 121–138. https://doi.org/10.1037/1082-989X.12.2.121.

Erath, K., Ingram, J., Moschkovich, J., & Prediger, S. (2021). Designing and enacting instruction that enhances language for mathematics learning: a review of the state of development and research. *ZDM—Mathematics Education*, *53*(2), 245–262. https://doi.org/10.1007/s11858-020-01213-2.

Fishman, E. J., Borko, H., Osborne, J., Gomez, F., Rafanelli, S., Reigh, E., & Berson, E. (2017). A practice-based professional development program to support scientific argumentation from evidence in the elementary classroom. *Journal of Science Teacher Education*, *28*(3), 222–249. https://doi.org/10.1080/1046560X.2017.1302727.

Florin, J. (2021). *Unterrichtsintegrierte Förderung des Operationsverständnisses*. Universität Basel. Dissertation

Gardner, R. (2001). *When listeners talk*. Benjamins.

Gellert, A. (2014). Students discussing mathematics in small-group interactions: Opportunities for discursive negotiation processes focused on contentious mathematical issues. *ZDM*, *46*, 855–869. https://doi.org/10.1007/s11858-014-0594-y.

Häsel-Weide, U., & Nührenbörger, M. (2021). Inklusive Praktiken im Mathematikunterricht. Empirische Analysen von Unterrichtsdiskursen in Einführungsphasen. *Zeitschrift für Grundschulforschung*, *14*(1), 49–65. https://doi.org/10.1007/s42278-020-00097-1.

Hiebert, J., & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade arithmetic. *American Educational Research Journal*, *30*(2), 393–425. https://doi.org/10.3102/00028312030002393.

Hiebert, J., Gallimore, R., Garnier, H., Bogard Givvin, K., Hollingsworth, H., Jacobs, J., & Stigler, J. (2003). *Teaching mathematics in seven countries*. U.S. Department of Education.

Howe, C., & Abedin, M. (2013). Classroom dialogue: a systematic review across four decades of research. *Cambridge Journal of Education*, *43*(3), 325–356. https://doi.org/10.1080/0305764X.2013.786024.

Howe, C., Hennessy, S., Mercer, N., Vrikki, M., & Wheatley, L. (2019). Teacher-student dialogue during classroom teaching: does it really impact on student outcomes? *Journal of the Learning Sciences*, *28*(4–5), 462–512. https://doi.org/10.1080/10508406.2019.1573730.

Ing, M., Webb, N. M., Franke, M. L., Turrou, A. C., Wong, J., Shin, N., & Fernandez, C. H. (2015). Student participation in elementary mathematics classrooms: the missing link between teacher practices and student achievement? *Educational Studies in Mathematics*, *90*(3), 341–356. https://doi.org/10.1007/s10649-015-9625-z.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: helping children learn mathematics*. National Academy Press.

Kocher, K., Krummenacher, R., Reusser, L., & Wagner, M. (2021). *Schweizer Zahlenbuch 2*. Klett & Balmer.

Krummheuer, G. (1997). *Narrativität und Lernen: mikrosoziologische Studien zur sozialen Konstitution schulischen Lernens*. Studien-Verlag.

Krummheuer, G. (2007). Argumentation and participation in the primary mathematics classroom: two episodes and related theoretical abductions. *The Journal of Mathematical Behavior*, *26*(1), 60–82. https://doi.org/10.1016/j.jmathb.2007.02.001.

Krummheuer, G. (2011). Representation of the notion "learning-as-participation" in everyday situations of mathematics classes. *ZDM—Mathematics Education,*, *43*(1), 81–90. https://doi.org/10.1007/s11858-010-0294-1.

Kultusministerkonferenz (2022). *Bildungsstandards im Fach Mathematik für den Primarbereich*. KMK.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310.

Mehan, H. (1979). *Learning lessons: social organization in the classroom*. Harvard University Press. https://doi.org/10.4159/harvard.9780674420106.

Michaels, S., & O'Connor, C. (2015). Conceptualizing talk moves as tools: professional development approaches for academically productive discussions. In L. B. Resnick, C. S. C. Asterhan & S. N. Clarke (Eds.), *Socializing intelligence through academic talk and dialogue* (pp. 347–361). AERA. https://doi.org/10.3102/978-0-935302-43-1_27.

Mok, S. Y., Hämmerle, C. S., Rüede, C., & Staub, F. S. (2022). How do professional development programs on comparing solution methods and classroom discourse affect students' achievement in mathematics? The mediating role of students' subject matter justifications. *Learning and Instruction*, *82*, 101668. https://doi.org/10.1016/j.learninstruc.2022.101668.

Moschkovich, J. N. (2007). Examining mathematical discourse practices. *For the Learning of Mathematics*, *27*(1), 24–30.

Moschkovich, J. N. (2015). Academic literacy in mathematics for english learners. *The Journal of Mathematical Behavior*, *40*(A), 43–62. https://doi.org/10.1016/j.jmathb.2015.01.005.

Moser Opitz, E., Stöckli, M., Grob, U., Reusser, L., & Nührenbörger, M. (2020). *BASIS-MATH-G 2+*. Hogrefe.

Murata, A., Siker, J., Kang, B., Baldinger, E. M., Kim, H. J., Scott, M., & Lanouette, K. (2017). Math talk and student strategy trajectories: the case of two first grade classrooms. *Cognition and Instruction*, *35*(4), 290–316. https://doi.org/10.1080/07370008.2017.1362408.

Nührenbörger, M., & Steinbring, H. (2009). Forms of mathematical interaction in different social settings: examples from students', teachers' and teacher-students' communication about mathematics. *Journal of Mathematics Teacher Education*, *12*(2), 111–132. https://doi.org/10.1007/s10857-009-9100-9.

O'Connor, C., Michaels, S., Chapin, S., & Harbaugh, A. G. (2017). The silent and the vocal: participation and learning in whole-class discussion. *Learning and Instruction*, *48*, 5–13. https://doi.org/10.1016/j.learninstruc.2016.11.003.

Pehmer, A.-K., Gröschner, A., & Seidel, T. (2015). Fostering and scaffolding student engagement in productive classroom discourse: Teachers' practice changes and reflections in light of teacher professional development. *Learning, Culture and Social Interaction*, *7*, 12–27. https://doi.org/10.1016/j.lcsi.2015.05.001.

Peterson, R. A., & Brown, S. P. (2005). On the use of beta coefficients in meta-analysis. *Journal of Applied Psychology*, *90*(1), 175–181. https://doi.org/10.1037/0021-9010.90.1.175.

Pfister, M., Moser Opitz, E., & Pauli, C. (2015). Scaffolding for mathematics teaching in inclusive primary classrooms: a video study. *ZDM*, *47*(7), 1079–1092. https://doi.org/10.1007/s11858-015-0713-4.

Praetorius, A. K., Pauli, C., Reusser, K., Rakoczy, K., & Klieme, E. (2014). One lesson is all you need? Stability of instructional quality across lessons. *Learning and Instruction*, *31*, 2–12. https://doi.org/10.1016/j.learninstruc.2013.12.002.

Praetorius, A. K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: the German framework of three basic dimensions. *ZDM—Mathematics Education*, *50*(3), 407–426. https://doi.org/10.1007/s11858-018-0918-4.

Resnick, L., Michaels, S., & O'Connor, C. (2010). How (well-structured) talk builds the mind. In D. D. Preiss & R. J. Sternberg (Eds.), *Innovations in educational psychology* (pp. 163–194). Springer.

Resnick, L. B., Asterhan, C. S. C., & Clarke, S. N. (Eds.). (2015). *Socializing intelligence through academic talk and dialogue*. AERA.

Royar, T. (2013). *Handlung – Vorstellung – Formalisierung: Entwicklung und Evaluation einer Aufgabenreihe zur Überprüfung des Operationsverständnisses für Regel- und Förderklassen*. Kovač.

Rüede, C., Mok, S. Y., & Staub, F. C. (2023). Fostering flexibility using comparing solution strategies and classroom discussion. Effects of two professional development programs. *Journal for Research in Mathematics Education*, *54*(1), 43–63.

Schindler, M. (2017). Inklusiven Mathematikunterricht gestalten lernen – praxisbezogen und zugleich handlungsentlastet. In J. Leuders, T. Leuders, S. Prediger & S. Ruwisch (Eds.), *Mit Heterogenität im Mathematikunterricht umgehen lernen* (pp. 199–209). Springer. https://doi.org/10.1007/978-3-658-16903-9_17.

Schnepel, S., Moser Opitz, E., Grob, U., Leuenberger, D., & Nührenbörger, M. *BASIS-MATH-G 1+*. Hogrefe. in preparation

Schütte, M., Jung, J., & Krummheuer, G. (2021). Diskurse als Ort der mathematischen Denkentwicklung – Eine interaktionistische Perspektive. *Journal für Mathematik-Didaktik*, *42*(2), 525–551. https://doi.org/10.1007/s13138-021-00183-6.

Schwarzkopf, R. (2015). Argumentationsprozesse im Mathematikunterricht der Grundschule: Ein Einblick. In A. Budke, M. Kuckuck, M. Meyer, F. Schäbitz, K. Schlüter & G. Weiss (Eds.), *Fachlich argumentieren lernen. Didaktische Forschungen zur Argumentation in den Unterrichtsfächern* (pp. 31–45). Waxmann. https://doi.org/10.25656/01:14021.

Sfard, A. (2008). *Thinking as communicating*. Cambridge University Press. https://doi.org/10.1017/CBO9780511499944.

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.

Steinbring, H. (2000). Mathematische Bedeutung als soziale Konstruktion im Mathematikunterricht. *Journal für Mathematikdidaktik*, *21*(1), 28–49. https://doi.org/10.1007/BF03338905.

Streit, C., & Rüede, C. (2023). *Operationsverständnis im Unterricht fördern – Ergebnisse einer Interventionsstudie*. submitted

Tabach, M., Hershkowitz, R., Azmon, S., & Dreyfus, T. (2020). Following the traces of teachers' talk-moves in their students' verbal and written responses. *International Journal of Science and Mathematics Education*, *18*(3), 509–528. https://doi.org/10.1007/s10763-019-09969-0.

Thanheiser, E., & Sugimoto, A. (2022). Justification in the context of elementary grades: justification to develop and provide access to mathematical reasoning. In K.N. Bieda, A. Conner, K.W. Kosko & M. Staples (Eds.), *Conceptions and consequences of mathematical argumentation, justification, and proof* (pp. 35–48). Springer. https://doi.org/10.1007/978-3-030-80008-6_4.

van der Veen, C., de Mey, L., van Kruistum, C., & van Oers, B. (2017). The effect of productive classroom talk and metacommunication on young children's oral communicative competence and subject matter knowledge: An intervention study in early childhood education. *Learning and Instruction*, *48*, 14–22. https://doi.org/10.1016/j.learninstruc.2016.06.001.

van der Veen, C., Michaels, S., Dobber, M., van Kruistum, C., & van Oers, B. (2021). Design, implementation, and evaluation of dialogic classroom talk in early childhood education. *Learning, Culture and Social Interaction*, *29*, 100515. https://doi.org/10.1016/j.lcsi.2021.100515.

Webb, N.M., Franke, M.L., Ing, M., Wong, J., Fernandez, C.H., Shin, N., & Turrou, A.C. (2014). Engaging with others' mathematical ideas: Interrelationships among student participation, teachers' instructional practices, and learning. *International Journal of Educational Research*, *63*, 79–93. https://doi.org/10.1016/j.ijer.2013.02.001.

Whitebread, D., Bingham, S., Grau, V., Pasternak, P.D., & Sangster, C. (2007). Development of metacognition and self-regulated learning in young children: role of collaborative and peer-assisted learning. *Journal of Cognitive Education and Psychology*, *6*(3), 433–455. https://doi.org/10.1891/194589507787382043.

Wood, T. (1999). Creating a context for argument in mathematics class. *Journal for Research in Mathematics Education*, *30*(2), 171–191. https://doi.org/10.2307/749609.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education*, *27*(4), 458–477. https://doi.org/10.5951/jresematheduc.27.4.0458.