"Let us discuss math"; Effects of shiftproblem lessons on mathematical discussions and level raising in early algebra

Calor, S.M.; Dekker, R.; van Drie, J.P.; Zijlstra, B.J.H.; Volman, M.L.L.

DOI
10.1007/s13394-019-00278-x

Publication date
2020

Document Version
Final published version

Published in
Mathematics Education Research Journal

Mathematics Education Research Journal (2020) 32:743–763
https://doi.org/10.1007/s13394-019-00278-x

ORIGINAL ARTICLE

"Let us discuss math"; Effects of shift-problem lessons on mathematical discussions and level raising in early algebra

Sharon M. Calor[1] · Rijkje Dekker[1] · Jannet P. van Drie[1] · Bonne J. H. Zijlstra[1] · Monique L. L. Volman[1]

Received: 13 July 2018 / Revised: 18 June 2019 / Accepted: 30 June 2019 /
Published online: 25 July 2019

Abstract

We investigated whether early algebra lessons that explicitly aimed to elicit mathematical discussions (shift-problem lessons) invoke more and qualitatively better mathematical discussions and raise students' mathematical levels more than conventional lessons in a small group setting. A quasi-experimental study (pre- and post-test, control group) was conducted in 6 seventh-grade classes ($N = 160$). An analysis of the interaction processes of five student groups showed that more mathematical discussions occurred in the shift-problem condition. The quality of the mathematical discussions in the shift-problem condition was better compared to that in the conventional textbook condition, but there is still more room for improvement. A qualitative illustration of two typical mathematical discussions in the shift-problem condition are provided. Although students' mathematical levels were raised a fair amount in both conditions, no differences between conditions were found. We concluded that shift-problem lessons are powerful for eliciting mathematical discussions in seventh-grade shift-problem early algebra lessons.

Keywords Shift-problem lessons · Early algebra · Collaborative learning · Mathematical level raising · Mathematical discussions · Secondary education

✉ Sharon M. Calor
s.m.calor@uva.nl

[1] Research Institute of Child Development and Education, Faculty of Social and Behavioral Sciences, University of Amsterdam, PO Box 15780, 1001 NG Amsterdam, The Netherlands

Introduction

Textbooks are commonly used in mathematics classrooms in the Netherlands, as well as in other countries. For the most part, these textbooks lead students step-by-step through predefined solution processes for solving mathematical assignments (Lithner 2008; Mayer 2002). However, it has been shown that students who were engaged in such step-by-step solving processes display superficial and fragmented mathematical knowledge and superficial mathematical reasoning (Bergqvist et al. 2008). Several experimental studies have shown that it is possible to invoke learning processes that contribute significantly to students' ability to reason mathematically and that mathematical discussions in small groups can lead to mathematical level raising (Dekker and Elshout-Mohr 1998, 2004; Pijls and Dekker (2011); Pijls et al. 2007). A major drawback of these experimental interventions is that their implementation in everyday mathematics classrooms would require substantial curriculum changes (Stein et al. 2007).

Palha et al. (2013) addressed this problem by developing design principles for lessons that aim to raise students' mathematical levels, based on curriculum materials that are used currently in mathematical lessons. These design principles are inspired by the domain specific theory of Realistic Mathematics Education of Freudenthal (1991). Main design principles are learning goals aiming at a deeper understanding of mathematics and invoking mathematical discussions through collaborative learning and challenging assignments. The mathematical content unfolds from the mathematical discussions and all assignments require reasoning. The authors referred to these lessons as shift-problem lessons, indicating that the mathematical problems that are discussed in these lessons are designed with the specific goal of a shift to a deeper understanding of mathematics. An example of such a shift-problem lesson for integral calculus is: "Given the graph representing the velocity of a car against time and the formula for the velocity. The students are asked to provide formulas that describe the distance travelled starting at different times." (Palha et al. (2014), p. 1596). In a quasi-experimental study, Palha et al. (2014) compared these shift-problem lessons in which 11th-grade students worked collaboratively in small heterogeneous groups (different levels of mathematical knowledge) with a conventional condition, in which the students worked individually on regular assignments of the textbook on the same topic. The authors found a significant positive effect on students' knowledge of integral calculus for shift-problem lessons when compared to conventional lessons. Furthermore, they found that students in the shift-problem condition were able to reason about integral calculus assignments at higher levels. Although these outcomes are promising, it is yet unclear whether the design principles that underpin shift-problem lessons will result in similarly positive outcomes when used with a different topic and with younger students.

The aim of this study is to investigate the effects of shift-problem lessons on the topic of early algebra for seventh-grade students (12- and 13-year-old students). We chose the topic early algebra because students' difficulties with learning early algebra in secondary education often reoccur in international comparative studies such as the Trends in International Mathematics and Science Study (TIMSS) (Mullis et al. 2016) and the Programme for International Student Assessment (PISA) (OECD 2016). In a quasi-experimental design, we compared the amount and quality of mathematical discussions between these shift-problem lessons and conventional lessons during group work, as well as the gain in students' mathematical levels between both conditions.

Theoretical framework

Mathematical level raising for early algebra

According to Freudenthal (1991) and Van Hiele (1986), the learning of mathematics occurs in discrete steps, implying the existence of levels of mathematical thinking. Van Hiele (1986) described four levels: (1) visual level: the forms of mathematical objects are the object of study, (2) descriptive level: the properties of mathematical objects are the object of study, (3) theoretical level: relations between the properties of mathematical objects are the object of study, and (4) formal logical level: relations between theorems are the object of study. Freudenthal (1991) built upon this theory and states that the levels are more relative than discrete, meaning that level raising occurs every time a mathematical activity (performed at a lower level) consciously becomes the object of reflection (at a higher level). Freudenthal (1971) defines a mathematical activity as an activity of organizing subject matter, which can be matter from reality or mathematical matter, according to mathematical patterns or new ideas. For example, at a lower level, patterns of the number of tables and chairs in a table setting can be studied. At a higher level, these patterns themselves become the object of reflection when one tries to create a formula with which the number of chairs in a particular table setting can be calculated when the number of corresponding tables is given.

According to Freudenthal (1978), when students collaborate in a heterogeneous small group on one task, there will often be at least one student who experiences an "Aha moment" (jumping to a higher level) when understanding the subject matter. A typical higher level activity will follow for that student, such as reflecting on how he/she mastered the subject matter and explaining to other students in the group what he/she just learned. In other words, discussions between students in a heterogeneous group while performing mathematical activities (mathematical discussions) enhance mathematical level raising.

A reoccurring problem with attaining mathematical level raising when learning early algebra in secondary education is that students find it difficult to make a shift from studying patterns at a lower level to understanding formulae at a higher level (Kieran 1992; Sfard and Linchevski 1994; Van Stiphout et al. 2011). Janvier (1987) addressed this problem by defining algebra representations, situation, graphs, tables, and formulae as four ways to describe the relation between two variables in a formula. He calls the

Table 1 Janvier translation skills. All possible translation skills between representations situation, tables, graphs, and formulae are shown

To From	Situation	Tables	Graphs	Formulae
Situation		Measuring	Sketching	Modeling
Tables	Reading		Plotting	Fitting
Graphs	Interpretation	Reading off		Curve fitting
Formulae	Parameter recognition	Computing	Sketching	

Adapted from Janvier (1987)

ability to switch between the representations a translation skill. Table 1 shows these translation skills.

According to Janvier (1987), translation skills are best learned when they are taught in a pair-wise manner. For example, the translation skill modeling is learned best when students first learn how to construct a situation from a formula, followed by learning how to construct a formula given a situation, or vice versa. In this study, all translation skills involving the representation formulae are considered to contribute to mathematical level raising. These translation skills are as follows: parameter recognition (formulae to situation), computing (formulae to tables), sketching (formulae to graphs), modeling (situation to formulae), fitting (tables to formulae), and curve fitting (graphs to formulae).

Collaborative learning and group composition

Freudenthal (1991) advocated collaborative learning for attaining mathematical level raising. Following Kaendler et al. (2015), we define collaborative learning as: "Collaborative learning is the process of two or more students working together to find a joint solution to the group task at hand." (Kaendler et al. 2015, p. 506). The focus is on co-constructing knowledge together, while students depend on one another because of their unique knowledge and perspectives. In this way, it differs from cooperative learning in which group tasks are divided in subtasks that can be solved individually (Dillenbourg 1999). The quality of interaction between students is an important indicator for the effectiveness of collaborative learning (Van Boxtel et al. 2000; Van der Linden et al. 2000; Ing et al. 2015).

Invoking mathematical discussions through collaborative learning is also one of the main design principles of shift-problem lessons. Collaborative learning in classrooms has been researched for decades (Johnson and Johnson 2009) and is known to have a positive effect on students' learning outcomes (Kyndt et al. 2014). It has been successful in promoting learning in the mathematics classroom (Dekker and Elshout-Mohr 1998, 2004; Pijls et al. 2007; Webb 2009; Yackel et al. 1991). Several studies have shown that collaborative problem-solving skills as giving and receiving explanations about mathematical content may lead to restructuring, clarification, and repairing of students' own knowledge and learning new knowledge (Webb 2009). Although collaborative learning is meaningful, the results of the PISA assessment on collaborative learning (OECD 2017) showed that only 8% of the students of the OECD countries seem to have high-level collaborative problem-solving skills.

One of the factors that may affect students' learning during collaborative learning is the composition of the group. Several studies have been conducted over the years to study the effect of group composition on learning outcomes for mathematics in secondary education. In their meta-analysis, Lou et al. (1996) did not find differences in performance between heterogeneous and homogeneous groups. Hooper and Hannafin (1988) reported similar results but also showed that heterogeneous grouping is more beneficial for students with low mathematical abilities. Similar results have been found in more recent studies (Webb 2011; Wiedmann et al. 2012). For instance, Wiedmann et al. (2012) showed that groups needed at least one member with high mathematical ability. More importantly, they reported that heterogeneous groups generated the highest quality approaches and the widest variety in problem-solving skills when learning early algebra.

Mathematical discussions

The effects of collaboration on learning depends on the quality of the discussion (Van Boxtel et al. 2000; Van der Linden et al. 2000; Ing et al. 2015). In mathematical discussions, students reason about mathematical subjects and are challenged to reflect on mathematical structures and activities (Freudenthal 1991; Van Hiele 1986). These practices are supposed to contribute to mathematical level raising. The two main activities of mathematical level raising mentioned by Freudenthal (reflection and discussion) form the basis of the Process Model of Dekker and Elshout-Mohr (1998).

The Process Model was developed to study the quality of mathematical discussions. The model is meant to analyze discussions in small group of students who work on a mathematical task. Students work on the same mathematical task, each in their own way. The Process Model distinguishes three types of learning activities: key activities, regulating activities, and mental activities. Key activities are communicative activities that help students attain mathematical level raising. Key activities invoke reflection; thus, they activate mental activities. Four key activities can be distinguished: to tell/show one's work, to explain one's work, to justify one's work, and to reconstruct one's work. For example, key activity "explanation" leads to mathematical level raising, since a student who does the explaining has to think about his work and thus may fill in gaps in his existing knowledge or may even reconstruct his existing knowledge (Webb et al. 2002). Regulating activities are communicative activities that are meant to regulate key activities. Three regulating activities can be distinguished: ask to show one's work, ask to explain one's work and criticize another students' work. And finally, mental activities are activities that occur in students' minds. Five types of mental activities can be distinguished: becoming aware of one's work, thinking about one's work, thinking about another students' criticism, thinking about one's justification, criticizing one's work. Key activities and regulating activities can be observed very well and can therefore also be measured very well. Mental activities take place in students' mind and are therefore more difficult to measure.

Table 2 shows the mental and key activities of student B, regulated by student A.

In this study, a mathematical discussion is considered to be a qualitatively good discussion if it comprises all key activities, since key activities evoke reflection which in its turn evokes mathematical level raising. Shift-problem lessons are a way to elicit these key activities, in particular, the design principle: "reflection can be induced through mathematical discussions" (Palha et al. 2013, p. 148–149).

Table 2 The three types of activities of the Process Model, regulating, mental, and key activities performed by student A and student B

Regulating Activities	Mental Activities	Key Activities
A asks B to tell/show his work	B becomes aware of his work	B tells/shows his work
A asks B to explain his work	B thinks about his work	B explains his work
A criticizes B's work	B thinks about A's criticism	B justifies his work
	B thinks about his justification	
	B criticizes his work	B reconstructs his work

Aims and research question

The aim of the current study is to investigate whether shift-problem lessons result in more and better quality mathematical discussions and more mathematical level raising for early algebra in the seventh grade. The focus of this study is not on the teacher. For both teachers and students, working in group settings was common practice. The focus of this study lies on material that could assist teachers in this setting. In a quasi-experimental study, we compared the effects of shift-problem lessons with a condition in which students worked with textbooks in a small group setting.

The following research question guides this study:

Do shift-problem lessons for early algebra in seventh grade result in more and qualitatively better mathematical discussions and more mathematical level raising than working with conventional lessons in a small group setting?

We hypothesized that

- students in the shift-problem condition would be engaged in more and better quality mathematical discussions than students in the conventional textbook condition.
- students in the shift-problem condition would reach a higher mathematical level than the students in the conventional textbook condition.

Method

Design

We used a quasi-experimental research design (pre–post-test, control group) to investigate our research question.

Intervention

In both conditions, students worked during 5 weeks on a lesson series of 12 lessons of 60 min on the topic of early algebra. In the shift-problem condition, we replaced five of these lessons, that were suitable for adaption, with shift-problem lessons. The shift-problem lessons consisted of tasks that were close to, or adaptations of, the conventional textbook tasks according to the sequence of the tasks in the textbook (Moderne Wiskunde 1A 2012; Moderne Wiskunde 1B 2012). By doing so, we aimed to stay close to the teacher's curriculum. We designed these tasks according to the design principles of Palha et al. (2013).

These design principles are:

1) "the designer is guided by the learning goal of a deeper understanding of mathematics
2) mathematics has to start at a level that is experientially real to the students and
3) reflection can be induced through mathematical discussions" (Palha et al. 2013, p. 148–149)

In none of the conditions, teachers were given prior instructions on how to support the small groups of students. Teachers gave support in their usual way (i.e., teachers mainly gave content help or hints to individual students in the groups).

In the conventional textbook condition, students sat together in small heterogeneous groups of three or four students (as was normally the case) and were allowed to talk to each other but worked individually on regular assignments from the conventional mathematics textbook during all 12 lessons. Every lesson started with 15 min instruction by the teacher, followed by students working on the textbook assignments for 45 min.

In the shift-problem condition, students worked collaboratively during five lessons (lesson 2, 4, 6, 11, and 12) on shift-problem lessons. Every lesson started with 15 min introduction by the teacher, followed by students working on the shift-problem lessons for 45 min. During the remaining lessons, students sat together in the same small heterogeneous groups and worked on regular assignments from the textbook.

Shift-problem lesson for early algebra

The main learning goal in the shift-problem lessons is a deeper understanding of early algebra, in particular, formulae (design principle: "the designer is guided by the learning goal of a deeper understanding of mathematics" (Palha et al. 2013, p. 148–149)). All possible switches to and from representation formulae are associated with this learning goal. To accomplish this learning goal, the lesson series (shift-problem lessons plus conventional sections), as a whole, contained exercises testing all possible switches between the Janvier (1987) representations. In particular, the lesson series contained exercises involving all six switches to and from representation formulae. If, for example, a particular switch only occurred in one direction in the conventional task, we added the corresponding switch in the opposite direction in the shift-problem lesson. By adapting conventional tasks, we stay close to students' experiences with mathematics (design principle: "mathematics has to start at a level that is experientially real to the students" (Palha et al. 2013, p. 148–149)) as Palha did with Geometry exercises (Palha et al. 2013).

For example, let us consider a task on early algebra in the conventional textbook (see Fig. 1).

This conventional task already contains the translation skill modeling (switch from situation to formulae). We adapted the task by adding the translation skill parameter recognition (switch from formulae to situation) to the task, so the translation skills parameter recognition and modeling can be learned in a pair-wise manner, as was suggested by Janvier (1987). Tables and chairs that are part of the conventional task were also printed in color and cut into separate figures of tables and chairs, so students could work with concrete material. Figure 2 shows the adapted task.

Teachers in both conditions implemented the lessons as intended (gave support as they normally would). Implementation check was performed by the first author during and after lessons (check with teacher and watch videotape of lessons).

Participants

Participants were 160 students, aged between 12 and 15 years old ($M = 12.87$, $SD = 0.54$), from 6 seventh-grade classes ($M = 26.67$, $SD = 0.52$), and 6 teachers (5 males, 1

An owner of a restaurant has a basic table setting of two rectangle tables beside each other, see table setting *A*. With a larger party of people, he extends the 2 tables with extra tables.

a How many people can take place in this manner at 3 tables?

b The setting is extended to 5 tables. Make a sketch of the situation.

c How many chairs are there around 5 tables?

d How many chairs are there at each table, if you disregard the outer ends?

e Write the formula in words for calculating the total number of chairs.

The number of tables times … plus … equals to the number of chairs

f You can write the formula in words shorter with a chain of arrows. Finish the chain of arrows.

IN-number OUT-number

number of tables $\xrightarrow{\times \ldots}$ … $\xrightarrow{+ \ldots}$ number of chairs

Fig. 1 Conventional task on table settings

female) of one school in an ethnically diverse suburban neighborhood in Amsterdam, the Netherlands. The first author is a teacher at this school.

Twenty-seven percent of the students spoke at least one foreign language at home along with Dutch. Eight percent of these ethnic minority students did not speak Dutch at all at home. The students were given track advice from primary school ranging from vocational levels to pre-university levels. In contrast to conventional practice in the Netherlands, the school delays streaming of students for 2 years according to primary track advisement. Seventy percent of the participating students were given track advice for a vocational level and 30% of the students for pre-university level. Students of different backgrounds and of different track advisement were distributed evenly over the classes by the school. Similar to Palha et al. (2014), we focused on heterogeneous groups in this study. All students (both conditions) worked in small heterogeneous groups (34 groups of four students and eight groups of three students). Student groups were formed based on the results of the pre-test on mathematical level, measured with a mathematical knowledge pre-test on early algebra. Students were divided into three categories based on their performances: weak, average, and strong. Groups consisted of one weak student, two average students and one strong student (Webb et al. 1998). Ideally students worked in groups of four students, but eight groups of three students were formed due to restrictions in class size.

The teachers volunteered to participate in this study. Every teacher taught his/her own class. The teachers were matched into three pairs according to age and teaching experience (6, 13, and 33 years of teaching experience). For every pair, one teacher was assigned to the experimental condition and the other to the control condition (three teachers with 17 groups of four students and four groups of three students in each

A setting of tables and chairs corresponds to the following formula in words:

The number of tables times 3 plus 2 equals to the number of chairs

a Construct a setting of tables and chairs together that corresponds to this formula in words.

Draw or glue this setting on your poster.

b Create a new formula in words that also corresponds to your setting.

Fig. 2 Adapted task on table settings with concrete material

condition, two males and one female in the experimental condition and three males in the control condition). One of the teachers is the first author of this article.

Data collection/instruments/analysis

Mathematical discussions

In every class (both conditions), one group was randomly selected for videotaping. The data from one group in the conventional textbook condition were missing, resulting in a total of three groups in the shift-problem condition and two groups in the conventional textbook condition. Prior to data collection, every group was videotaped during one lesson for students to get used to the camera.

Group interaction was videotaped in lesson 3. Shift-problem lesson 3 was selected for its clear application of the design principles that underpin shift-problem lessons.

Interaction processes were transcribed and analyzed by using the program Multiple Episode Protocol Analysis (Erkens 2002). The interaction was coded on the level of utterances. Following Van Boxtel et al. (2000), we define an utterance as "an individual message unit that is distinguished from another utterance through a 'perceptible' pause, comma or full stop" (Van Boxtel et al. 2000, p. 317–318). To get a general impression of the interaction processes, we first coded all utterances on the level of task acts. Task acts refer to the function of the utterances in relation to the execution of the task (Van Drie et al. 2005). We distinguished seven main categories; four categories on the dimension of on-task utterances and three categories off-task utterances (see Table 3). In the second step, utterances that were coded as talk about task content were divided further into talk about task content to the teacher and mathematical discussions (talk about task content to other students in the group). Finally, in the third step, utterances coded as mathematical discussions in the second step were divided further into seven subcategories according to the regulating and key activities of the Process Model (Dekker and Elshout-Mohr 2004): ask to show work, ask to explain work, criticize work, tell/show work, explain work, justify work, and reconstruct work (see Table 4). In the "Two examples of mathematical discussions in the shift-problem condition" section, we provide two examples of the analyses of typical mathematical discussions in the shift-problem condition.

Inter-rater reliability between two coders, the first author and a research assistant, was calculated over three randomly selected protocols (sum of 1236 utterances) for all three coding steps. Inter-rater reliability was good (agreement 82% and Cohen's kappa 0.76) for the coding in the first step (talk about task content, think aloud about task content, talk about the task in general, talk about performing the task, social talk, talk about the camera, no code). Inter-rater reliability was excellent (agreement 99.7% and Cohen's Kappa 0.97) for the second coding step after agreeing on the first step (talk about task content with the teacher and elements of mathematical discussions), and inter-rater reliability was satisfactory (agreement 84% and Cohen's kappa 0.67) for the third coding step after agreeing on the second step (detailed analysis of mathematical discussions in regulating and key activities). For the third step, Cohen's Kappa turned out to be satisfactory but lower than may be expected due to an artifact in the calculation of Cohen's Kappa. Because subcategory tell/show work occurs the most and others occur scarcely, the distribution of the diagonal of the matrix used in the

Table 3 Coding definitions of task acts

Main categories	Description	Example in this study
On task		
Talk about task content		
Talk with teacher	Conversations with teacher about mathematical content	"Miss, can you explain to us why it should be doubled?"
Mathematical discussions	Conversations between students in a small group about mathematical content	"I have 6, because you have to double it"
Think aloud about task content	Think aloud about mathematical content, no interaction	"5 times 3 is umm … 15"
Talk about the task in general	Read the exercise out loud	"Can you create a formula in words?"
	Evaluation of the task	"This task was difficult"
	Cite what should be written down	"Write down: In the candy store…"
	Cite what the answer is	"The answer was 12"
	Conversations about the material	"Can you hand me the chairs?"
Talk about performing the task	Encourage other students to work	"You should also help"
	Planning task	"We have 5 min for the last exercise"
Off Task		
Social talk	Non-task related conversations with other students and teacher	"What subject do we have next?" "Teacher, when will we get our grades?"
Talk about the camera	Non-task related conversations about the camera and microphone	"When you tap on the microphone, you hear this"
No code	Inaudible	"[inaudible]"

calculation of Cohen's Kappa is uneven, leading up to a lower value for Cohen's Kappa than otherwise would be the case (Birt et al. 1993). The research assistant independently coded the data of two groups in the experimental condition, and the first author coded the rest of the data.

Table 4 Seven elements of a mathematical discussion

Regulating activities	Example
Ask to show work	"What do you have?"
Ask to explain work	"How did you do that?"
Criticize work	"That is in the wrong order"
Key activities	Example
Tell/Show work	"Look what I have done"
Explain work	"I have 6, because you have to double it"
Justify work	"I am right, because 5 times 5 = 25"
Reconstruct work	"I'll better do it like this"

We compared absolute frequencies of the categorical elements of mathematical discussions to determine if more elements of mathematical discussions occur in the shift-problem condition than in the conventional textbook condition.

Mathematical level

Students' mathematical levels in early algebra were measured by means of a test. A pre-test was administered in the lesson prior to the intervention, and the same test was administered as a post-test in the lesson after the intervention. The test aims to measure the mathematical levels of students based on the translation skills from Janvier (1987). The highest level implies being able to translate from representation formulae to representations situation, tables, and graphs and vice versa.

The test consists of six questions and 17 sub-questions, of which 10 were used to measure mathematical levels. They consist of questions involving all translation skills of the representation formulae (Janvier 1987). The other seven questions were discarded (not included in the scores) during determination of level raising. They consist of basic primary school arithmetic questions so that students would be able to answer at least some of the questions when the test was administered as a pre-test. Scores of 0, 1, and 2 were assigned; 0 meaning low mathematical level, 1 meaning medium mathematical level, and 2 meaning high mathematical level. The sub-questions of the test corresponded to the switches between the Janvier representations involving formulae. The first sub-question only involved a less abstract representation of a formula; therefore, the score for this question was limited to 1 (medium mathematical level). The sum of the scores amounts to a maximal number of 19 points in total that students could score for the test. The levels that were assigned to the sub-questions were discussed with a second coder (the second author of this article). Students worked individually on the test for 60 min. An example of a basic primary school arithmetic question is: "Out of one package pancake mix you can bake six pancakes. How many pancakes can you bake with 12 packages of pancake mix?" An example of a question that measures mathematical level is: "Shane wants to get in shape, therefore he goes swimming. With the following (word) formula you can calculate Shane's costs. Number of times swimming $\times 3 + 30 =$ costs. (Here, 3 stands for €3, the cost for swimming once at a swimming pool for a member and 30 stands for €30, the cost of a yearly membership to a swimming pool.) Draw a graph for this formula."

Inter-rater reliability for the pre- and post-test between two coders (the first and second author) over 36 randomly selected tests (10% of the total number of tests) was excellent (ICC 0.95). The first author coded the rest of the data.

A multilevel model for repeated observations on fixed occasions with an unrestricted covariance matrix (Snijders and Bosker 2012) was used to test for differential growth between the shift-problem condition and the conventional textbook condition. The fixed occasions are the measurements (pre- and post-test) nested in individual students. The measurements (pre- and post-test) are the first level and the individual students are the second level.

Results

In this section, we first show the results on the interaction processes during mathematical discussions in section 5.1, followed by two examples of such mathematical discussions in section 5.2, and finally, we show the results on the mathematical level test in section 5.3 to explore the research question, "Does working with Shift-Problem Lessons for Early Algebra in seventh grade lead to more mathematical discussions and more mathematical level raising than working with conventional collaborative lessons?"

Interaction processes during mathematical discussions

Table 5 shows absolute frequencies of the main categories of the interaction analysis of group discussions. The average number of total utterances in the Shift-problem condition (SPC) ($M = 596$) is approximately equal to the average number of total utterances in the Conventional textbook condition (CC) ($M = 592.5$). However, in the SPC, there are more utterances of the categories of mathematical discussions, talk about task in general, and talk about performing the task than in the CC. Utterances of the category Think aloud about task sparsely occur in both conditions. There are less utterances of category Social talk in the SPC than in the CC.

In Table 6, we zoom in on the elements of the mathematical discussions. As mentioned earlier, more utterances of the category of mathematical discussions occur in the SPC than in the CC. In groups 1, 2, and 3 of the SPC, the distribution of activities is relatively the same. Regulating activities ask to show work and criticize work occur the most while ask to explain work sparsely occurs. Of the key activities, tell/show work occurs the most while explain work, justify work, and reconstruct work sparsely occur. In group 1, explain work does not even occur at all. In the CC, the regulating

Table 5 Number of utterances and percentages per main interaction category summed over shift-problem lesson 3 for three groups in the shift-problem condition and two groups in the conventional textbook condition

Shift-problem lesson 3	Shift-problem condition			Conventional textbook condition	
	Group 1	Group 2	Group 3	Group 4	Group 5
Total utterances	737	693	358	620	565
On task					
Talk about task content					
Talk with teacher	14	5	10	6	3
Mathematical discussions	195	182	68	30	0
Think aloud about task	2	1	0	3	0
Talk about task in general	171	157	67	137	28
Talk about performing the task	51	83	24	5	23
Off task					
Social talk	265	241	153	419	495
Talk about camera	26	23	19	6	15
No code	13	1	17	14	1

Table 6 Elements of mathematical discussions summed over shift-problem lesson 3 for three groups in the shift-problem condition and two groups in the conventional textbook condition

Shift-problem lesson 3	Shift-problem condition			Conventional textbook condition	
	Group 1	Group 2	Group 3	Group 4	Group 5
Total elements of mathematical discussions	195	182	68	30	0
Regulating activities					
Ask to show work	26	24	6	9	0
Ask to explain work	2	0	0	0	0
Criticize work	33	29	14	0	0
Key activities					
Tell/show work	114	115	45	20	0
Explain work	0	6	1	1	0
Justify work	15	5	2	0	0
Reconstruct work	5	3	0	0	0

activity ask to show work occurs the most in group 4 while ask to explain work and criticize work do not occur at all. The key activity tell/show work also occurs the most and explain work once while justify work and reconstruct work do not occur at all. In group 5, no mathematical discussions at all occur between the students. It is clear that the key activity tell/show work occurs the most in both conditions while other key activities sparsely occur.

Two examples of mathematical discussions in the shift-problem condition

We provide two examples of the analyses of the mathematical discussions of a group of students (who we address with fictitious names). In the first example, the group of

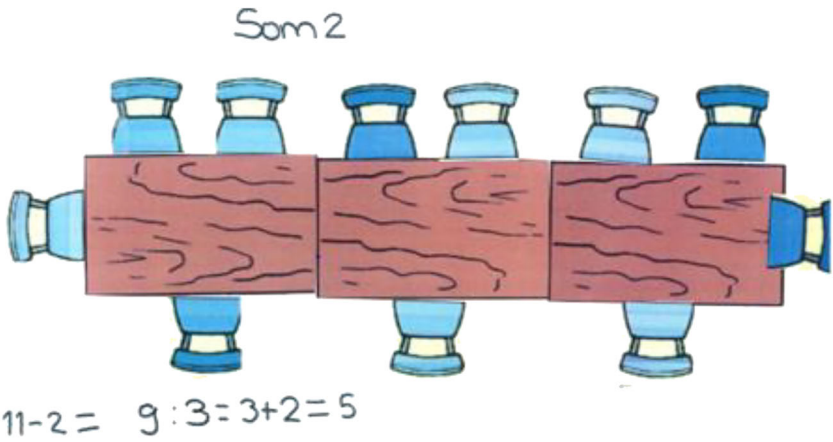

Fig. 3 Constructed table setting from students Leo, Hanna, Elizabeth, and Farrah corresponding to the formula in the following words: *the number of tables times 3 plus 2 equals the number of chairs*

students, Leo (strong student), Elizabeth (average student), Farrah (average student), and Hanna (weak student), do not manage to co-construct the knowledge they need to solve the task independently of the teacher. In the second example, the group does manage to co-construct the knowledge they need to collaboratively solve the task.

In example 1, the group worked on the adapted task in Fig. 2. First, students were asked to construct a table setting corresponding to the formula in words: "The number of tables times 3 plus 2 equals to the number of chairs" (sub-question a). Second, students were asked to create a new formula that also corresponds to their table setting (sub-question b).

Figure 3 shows the table setting the students constructed for the formula in words while working on sub-question a of the adapted task in Fig. 2. Students had no difficulty with the creation of this table setting.

However, students did have difficulty with the creation of a formula that also corresponds with the table setting they created (sub-question b). One way to solve the task is to reverse the formula in words by starting with the number of chairs instead of the number of tables. The new formula then becomes: the number of chairs minus 2 divided by 3 equals the number of tables. In the exercise, the hint was given that the formula could start with "number of chairs".

In the following coded excerpt, the students discussed the creation of this new formula, but were not able to co-construct the knowledge to create this new formula independently. It represents a typical mathematical discussion with much of the key activity tell/show work.

Leo:	Create a new formula in words that also corresponds to this setting (reads)	
Leo:	It has to start with number of chairs	
Elizabeth:	Easy	
Leo:	Number of chairs	Tell/show work
Hanna:	Number of chairs, no, number of chairs times 3, plus 2	Tell/show work
Leo:	Times 3 plus 2 (simultaneous with Hanna)	Tell/show work
Elizabeth:	Wait! How many chairs are there? 4, 5, 6, 7, 8, 9, 10	Criticize work
Leo:	2, 4, 6, 7, 8, 9, 10	Tell/show work
Elizabeth:	So no	Criticize work
Leo:	9, 10, 11	Tell/show work
Hanna:	11 times 3	Tell/show work
Leo:	And plus 2	Tell/show work
Hanna:	Plus 2. Equals to 35	Tell/show work
Leo:	But there aren't 35 chairs	Criticize work
Farrah:	3, 3, 3 is 9	Tell/show work
Hanna:	Yes, but	
Farrah:	10, 11. Is 11 chairs	Tell/show work
Elizabeth:	Okay	Tell/show work
Leo:	Yes, but what is the formula in words?	Ask to show work
Hanna:	O, number of tables times	Tell/show work
Elizabeth:	What is the formula in words, then?	Ask to show work
Leo:	I really don't know. Ask the teacher? Teacher, teacher?	

	0	1	2	5
	25	31	37	55

a Create a formula that corresponds to the table above.

b Create a story that corresponds to the table and your formula.

Fig. 4 Adapted task create a formula and story from a table

The discussion starts with general talk about the task. Leo reads the question and hints out loud, and Elizabeth evaluates the question as being easy. Leo then shows his work by following the hint that the new formula could begin with "number of chairs." Hanna shows her work by filling in the rest of the formula, number of chairs times 3 plus 2, simultaneously with Leo. This formula is incorrect (it might be just a repetition of the second part of the original formula). Elizabeth then criticizes Leo and Hanna's work by asking how many chairs are there, thus implying that the corresponding number of tables in the table setting does not fit Leo and Hanna's formula. Leo answers 11 and Hanna calculates that according to their formula, there should be 35 chairs in the table setting (both tell/show work). Leo then criticizes his and Hanna's formula by stating that there are not 35 chairs in the table setting, so the formula could not possibly be correct. Farrah then joins the discussion after having counted the number of chairs in the table setting and tells/shows her work by telling that there are 11 chairs. Leo and Elizabeth then ask the other students what the formula could be (ask to show work). None of the students knows the answer, so Leo suggests asking the teacher for help.

In the second example, the group worked on the first part (sub-question a) of an adapted task in which they were asked to create a formula that corresponds to the given table (see Fig. 4).

The group was able to co-construct the knowledge to solve the first part of the task collaboratively.

In the following coded excerpt, the students discussed the creation of their formula. It shows how students collaboratively solve the task. In contrast to the previous example, this example contains all four key activities, tell/show work, explain work, justify work, and reconstruct work, but still key activity tell/show work occurs the most.

Leo:	Create a formula that corresponds to the table above. (reads)	
Leo:	A formula, is it not number of times ×	Tell/show work
Elizabeth:	No, it is number of times	Criticize work
Elizabeth:	But wait, 31 minus 25, let me calculate for a minute	Tell/show work
Leo:	Why? Why do you do that?	Ask to explain work
Elizabeth:	Oh no, wait, 25, 25, … (louder)	Reconstructs work
Farrah:	I know it! Number of times ×...	Tell/show work
Leo:	Take it easy (to Elizabeth), it is number of times …	Tell/show work
Farrah:	Number of times × 1 plus 25	Tell/show work
Leo:	Why?	Ask to explain

Hanna:	What? (skeptical)	Criticize work
Hanna:	Isn't it 25 times 1 plus …	Tell/show work
Elizabeth:	In any case 25 (simultaneously with Farrah)	Tell/show work
Farrah:	In any case 25 (simultaneously with Elizabeth)	Tell/show work
Elizabeth:	That in any case	Tell/show work
Farrah:	Yes	Tell/show work
Leo:	Number of times ×	Tell/show work
Farrah:	So, number of times × something plus 25 or so	Tell/show work
Leo:	But it does not have to contain a plus all the time?	Ask to explain
Elizabeth:	Yes, that has to be the case	Tell/show work
Farrah:	Because, look, look, 25. It starts with 25 and then times, 25 times	Explain work
Hanna:	No!	Criticize work
Farrah:	Oh, wait times 6	Reconstructs work
Leo:	But, than it is the same. Number of times × 5 plus 20 (points to the graph of the previous exercise).	Tell/show work
Farrah:	No	Criticize work
Farrah:	Number of times ×6 plus 25	Tell/show work
Leo:	Why is it 6?	Ask to explain
Farrah:	Look, in between here is 6, 6, 6 (points to the items in the table)	Explain work
Elizabeth:	Do you understand?	
Leo:	Ooooh	
Farrah:	Number of times × 6 plus 25 (dictates to Leo)	
Elizabeth:	That's why I subtracted it. That's why I did minus	Justify work
Leo:	Number of times × 6 plus 25 (writes on poster)	

The discussion starts with general talk about the task. Leo reads the question out loud and begins showing his work by suggesting that the formula might be something like number of times ×. Elizabeth immediately criticizes Leo's work without letting him finish his sentence. She shows her work by suggesting that the solution might start with number of times ×, but thinks of another approach to solve the problem. She shows her work by elaborating on this approach, 31 minus 25. Elizabeth has weak arithmetic skills, so she needs some time to make this calculation. In the meanwhile, Leo asks Elizabeth to explain why she subtracts 25 from 31. Elizabeth might think that Leo's question is some kind of criticism, because she begins to doubt her approach and reconstructs her work into something beginning with 25. Then, Farrah shows her work by suggesting number of times × 1 plus 25. This formula is incorrect. Leo asks Farrah to explain her formula. Farrah ignores him. Hanna, who has been listening throughout the whole discussion, is skeptical about Farrah's formula and criticizes Farrah's work. She shows her own work by offering her formula, 25 times 1 plus … as a solution. This formula is also incorrect. Elizabeth and Farrah consequently co-construct the knowledge that, in any case, the formula has to contain 25 in it. Leo then shows his work, number of times ×, followed by Farrah who almost has the solution and shows her work, number of times × something plus 25 or so. Leo asks her to explain whether or not the formula should always contain a plus in it. Elizabeth shows her work by answering that that is indeed always the case. Then Farrah

starts to explain to Leo why this is the case, by explaining that the table starts with 25, meaning that 25 has to be multiplied (which is incorrect). Hannah then criticizes Farrah's explanation. Farrah suddenly understands that times 6 is the correct answer and seems to experience what Freudenthal called an "Aha moment" (jumping to a higher level) (Freudenthal 1978), 'Oh wait!'. She reconstructs her work, using times 6. Leo then shows his work by saying that it is the same formula as that of the previous exercise (which is incorrect). Farrah denies that and shows her work by naming the whole correct formula, number of times × 6 plus 25. Leo asks her to explain why it is times 6. Farrah explains to him why that is the case. Leo understands the explanation. Elizabeth realizes that her earlier approach to solve the task was correct and justifies why she subtracted 25 from 31 in the beginning. Farrah dictates the formula to Leo, who writes it down.

Mathematical level raising

We present the means of the pre- and post-test on mathematical level for the shift-problem condition and conventional textbook condition in Table 7.

Table 7 shows that the students' mathematical level improved a fair amount in both conditions. For the mathematical condition, the mean score increased from 1.43 to 9.53, and for the CC, it increased from 1.72 to 8.80.

A multilevel model for repeated observations on fixed occasions with an unrestricted covariance matrix (Snijders and Bosker 2012) was used to test for differential growth between the SPC and the CC.

We report on the model for two levels (with measurements (pre-test post-test) nested in individual students) since there was no model improvement for three levels (measurements, individuals, and groups) and four levels (measurements, individuals, groups, classes).

The expected outcome for the CC was pre-test = 1.42 and post-test = 1.42 + 9.5136. For the SPC, the expected outcome was pre-test = 1.42 + 0.30 and post-test = 1.42 + 0.30 + 9.5136 – 1.01. Thus, compared to CC the growth in the SPC is 1.01 smaller. This differential growth is not significant ($p = 0.19$).

Discussion/conclusion

The aim of this study was to investigate whether shift-problem lessons, lessons in which the design principles that underpin shift-problem lessons are applied for the topic

Table 7 Means and standard deviations pre- and post-test on mathematical level, shift-problem condition ($N = 180$) and conventional textbook condition ($N = 180$), maximum score of 19

	Shift-problem condition *M (SD)*	Conventional textbook condition *M (SD)*
Pre-test	1.43 (2.28)	1.72 (2.40)
Post-test	9.53 (4.69)	8.80 (4.86)

of early algebra, would result in more and qualitatively better mathematical discussions and more mathematical level raising than conventional collaborative lessons.

According to Freudenthal (1991), discussion and reflection are the main activities in mathematical level raising. We expected that more mathematical discussions, in which students are challenged to reflect on mathematical structures and activities, would occur in the shift-problem condition and that more mathematical level raising would occur in the shift-problem condition than in the conventional textbook condition.

First, our study showed that shift-problem lessons on the topic of early algebra in the seventh-grade students elicited more and qualitatively better mathematical discussions than conventional textbook lessons. Whereas mathematical discussion did occur in the shift-problem lessons, it was hardly or not found in the conventional textbook lessons.

In addition, we examined the quality of the discussion. A mathematical discussion is considered to be of good quality if it consists of all key activities, tell/show work, explain work, justify work, and reconstruct work. The key activity that occurred most in the discussions was tell/show work. Furthermore, we found that the discussions in the shift-problem condition were of better quality compared to the conventional textbook condition. Other key activities were found in the shift-problem condition, such as justify work and reconstruct work. Additionally, more regulating activities were found, such as ask to show work and criticize work. For the one group in the conventional textbook conditions in which mathematical discussions did occur, only ask to show work and one time explain work was found. It thus can be concluded that the shift-problem lessons did indeed elicit more mathematical discussions and that these discussions were of better quality. Still, improvement is necessary as most activities were telling and showing work. To gain a deeper understanding more explaining, justifying and reconstruction activities should occur (Dekker and Elshout-Mohr 1998).

With respect to our second hypothesis, we found that students' mathematical levels were raised in both conditions. However, multilevel analyses did not show differences between the conditions.

A possible explanation for not finding more mathematical level raising in the shift-problem lessons might be found in the quality of the discussions. As indicated before, the quality of the discussion is related to learning (Van der Linden et al. 2000). Key activities as formulated by Dekker and Elshout-Mohr (1998) are associated with reflection, which is related to mathematical level raising (Freudenthal 1991). Although the quality of the mathematical discussion was better in the shift-problem condition, as more and more diverse activities occurred, the fact that the key activity tell/show work occurred by far the most might explain why students did not achieve more level raising in the shift-problem condition.

Our findings are not in line with the finding of Palha et al. (2014). A first difference between the two studies is the topic, early algebra versus integral calculus. However, in our experience, the design principles were well suited to design lessons for early algebra. Second, the teachers in Palha's study had a higher teacher degree, which is required to be able to teach at pre university level in the Netherlands, than the teachers in our study. Third, students in the experiment of Palha et al. (2014) were older, had more knowledge of mathematics, and had chosen a profile with a difficult variant of mathematics. More research is needed to determine what the effects of shift-problem lessons are for different topics and different levels and ages of students.

This study was conducted at one school, which has limitations for the generalization of our findings. Still, several classes and several teachers participated in this study. This research could be replicated on a larger scale to find more robust findings. For the analysis of the interaction processes, a limited number of group discussions were randomly selected for analysis. Weaker or stronger groups could have been selected for analysis, which might have influenced our findings. It might thus be interesting to analyze discussions of other groups and of other lessons.

Our coding scheme enabled us to analyze in detail students' conversations during group work and most importantly the quality of the mathematical discussions. This analysis sheds light on the learning processes that occurred. Still, it is difficult to capture all learning processes that occur. For example, the key activity reconstruct work also might occur in students' minds, which cannot be measured.

To conclude, this research suggests that shift-problem lessons do invoke more mathematical discussions than conventional lessons. However, we have to be cautious in generalizing these results, as they were collected at only one school. This research also shows that applying shift-problem lessons is a necessary but not sufficient condition to achieve more mathematical level raising. The quality of the mathematical discussions in the shift-problem condition might be the key to increasing mathematical level raising, as the key activity tell/show work occurred the most, and others hardly occurred. This outcome raises a question regarding how teachers could improve the quality of the mathematical discussion and, hopefully, mathematical level raising. One possible way might be through scaffolding (Van de Pol et al. 2012). Future research should focus on the supporting role of the teacher in stimulating mathematical level raising.

References

Bergqvist, T., Lithner, J., & Sumpter, L. (2008). Upper secondary students' task reasoning. *International Journal of Mathematical Education in Science and Technology, 39*(1), 1–12.

Birt, T., Bishop, J., & Carlin, J. B. (1993). Bias, prevalence and kappa. *Journal of Clinical Epidemiology, 46*, 423–429.

Dekker, R., & Elshout-Mohr, M. (1998). A process model for interaction and mathematical level raising. *Educational Studies in Mathematics, 35*(3), 303–314.

Dekker, R., & Elshout-Mohr, M. (2004). Teacher interventions aimed at mathematical level raising during collaborative learning. *Educational Studies in Mathematics, 56*(1), 39–65.

Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.), *Collaborative-learning: cognitive and computational approaches* (pp. 1–19). Oxford, UK: Elsevier.

Erkens, G. (2002). Introduction to MEPA. Retrieved from http://edugate.fss.uu.nl/mepa Accessed 8 March 2014.

Freudenthal, H. (1971). Geometry between the devil and the deep sea. *Educational Studies in Mathematics, 3*(3), 413–435.

Freudenthal, H. (1978). *Weeding and sowing: preface to a science of mathematics education.* Dordrecht, the Netherlands: Reidel.

Freudenthal, H. (1991). *Revisiting mathematics education*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Hooper, S., & Hannafin, M. (1988). Cooperative CBI: the effects of heterogeneous versus homogeneous grouping on the learning of progressively complex concepts. *Journal of Educational Computing Research, 4*, 413–424.

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.

Janvier, C. (1987). Translation process in mathematics education. In C. Janvier (Ed.), *Problems of representation in mathematics learning and problem solving* (pp. 27–31). Hillsdale, NJ: Lawrence Erlbaum Associates.

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: social interdependence theory and cooperative learning. *Educational Researcher, 38*, 365–379.

Kaendler, C., Wiedmann, M., Rummel, N., & Spada, H. (2015). Teacher competencies for the implementation of collaborative learning in the classroom: a framework and research review. *Educational Psychology Review, 27*(3), 505–536. https://doi.org/10.1007/s10648-014-9288-9.

Kieran, C. (1992). The learning and teaching of school algebra. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 390–419). NewYork, NY: Macmillan.

Kyndt, E., Raes, E., Lismont, B., Timmers, F., Dochy, F., & Cascallar, E. (2014). A meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings? *Educational Research Review, 10*, 133–149.

Lithner, J. (2008). A research framework for creative and imitative reasoning. *Educational Studies in Mathematics, 67*(3), 255–276.

Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & Apollonia, S. (1996). Within-class grouping: a meta-analysis. *Review of Educational Research, 66*(4), 423–458.

Mayer, R. E. (2002). Rote versus meaningful learning. *Theory Into Practice, 41*(4), 226–232.

Moderne Wiskunde 1A vmbo-gt(h) (10th ed.) (ISBN: 978 9001 81378 9). (2012). Groningen, the Netherlands: Noordhoff Uitgevers bv.

Moderne Wiskunde 1B vmbo-gt(h) (10th ed.) (ISBN: 978 9001 81381 1). (2012). Groningen, the Netherlands: Noordhoff Uitgevers bv.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *Timms2015 International Results in Mathematics* (ISBN: 978 1 889938 29 5). Retrieved from http://timssandpirls.bc.edu/timss2015/international-results/timss-2015/mathematics/achievement-in-content-and-cognitive-domains/achievement-in-mathematics-content-domains/.

OECD. (2016). *PISA 2015 results (volume I): excellence and equity in education*. Paris, France: OECD Publishing. https://doi.org/10.1787/9789264266490-en.

OECD. (2017). *PISA 2015 results (volume V): collaborative problem solving*. Paris, France: OECD Publishing. https://doi.org/10.1787/9789264285521-en.

Palha, S., Dekker, R., Gravemeijer, K., & Van Hout-Wolters, B. (2013). Developing shift problems to foster geometrical proof and understanding. *The Journal of Mathematical Behavior, 32*(2), 142–159.

Palha, S., Dekker, R., & Gravemeijer, K. (2014). The effect of shift-problem lessons in the mathematics classroom. *International Journal of Science and Mathematics Education., 13*, 1589–1623. https://doi.org/10.1007/s10763-014-9543-z.

Pijls, M., & Dekker, R. (2011). Students discussing their mathematical ideas: the role of the teacher. *Mathematics Education Research Journal, 23*(4), 379–396. https://doi.org/10.1007/s13394-011-0022-3.

Pijls, M., Dekker, R., & Van Hout-Wolters, B. (2007). Teacher help for conceptual level raising in mathematics. *Learning Environments Research, 10*(3), 223–240.

Sfard, A., & Linchevski, L. (1994). The gains and pitfalls of reification – the case of algebra. *Educational Studies in Mathematics, 26*, 191–228.

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: an introduction to basic and advanced multilevel modeling* (second ed.). London, England: Sage Publishers.

Stein, M.K., Remmilard, J., & Smith, M.S. (2007). How curriculum influences student learning. In F.K. Lester(Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 319-370). Charlotte, NC: Information Age Publishing.

Van Boxtel, C., Van der Linden, J., & Kanselaar, G. (2000). Collaborative learning tasks and the elaboration of conceptual knowledge. *Learning and Instruction, 10*(4), 311–330.

Van de Pol, J., Volman, M., & Beishuizen, J. (2012). Promoting teacher scaffolding in small-group work: a contingency perspective. *Teaching and Teacher Education, 28*(2), 193–205.

Van der Linden, J. L., Erkens, G., Schmidt, H., & Renshaw, P. (2000). Collaborative learning. In R. J. Simons, J. L. van der Linden, & T. Duffy (Eds.), *New Learning* (pp. 37-55). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Van Drie, J., Van Boxtel, C., Jaspers, J., & Kanselaar, G. (2005). Effects of representational guidance on domain specific reasoning in CSCL. *Computers in Human Behavior, 21*(4), 575–602.

Van Hiele, P. M. (1986). *Structure and insight*. Orlando, FL: Academic Press.

Van Stiphout, I., Drijvers, P., & Gravemeijer, K. (2011). The development of students' algebraic proficiency. *International Electronic Journal of Mathematics Education, 8*(2), 62–80.

Webb, N. M. (2009). The teacher's role in promoting collaborative dialogue in the classroom. *British Journal of Educational Psychology, 79*, 1–28.

Webb, N. M. (2011). An analysis of group interaction and mathematical error in heterogeneous ability groups. *British Journal of Educational Psychology., 50*, 266–276. https://doi.org/10.1111/j.2044-8279.1980.tb00810.x.

Webb, N. M., Nemer, K. M., Chizhik, A. W., & Sugrue, B. (1998). Equity issues in collaborative group assessment: group composition and performance. *American Educational Research Journal, 35*(4), 607–651.

Webb, N. M., Nemer, K. M., & Zuniga, S. (2002). Short circuits or superconductors? Effects of group composition on high-achieving students' science assessment performance. *American Educational Research Journal, 39*(4), 943–989.

Wiedmann, M., Leach, R., Rummel, N., & Wiley, J. (2012). Does group composition affect learning by invention? *Instructional Science, 40*, 771–740. https://doi.org/10.1007/s11251-012-9204-y.

Yackel, E., Cobb, P., & Wood, T. (1991). Small-group interactions as a source of learning opportunities in second-grade mathematics. *Journal for Research in Mathematics Education, 22*(5), 390–408.