The handle http://hdl.handle.net/1887/67090 holds various files of this Leiden University dissertation.

Author: Houwelingen, M.J. van
Title: Measurement numeracy education for prospective elementary school teachers: effects of inductive and deductive teaching on classroom interaction and student performance
Issue Date: 2018-11-27
CHAPTER 7

General discussion

The previous chapters of this dissertation report on differential effects of two didactical approaches (inductive and deductive) to measurement numeracy education for prospective elementary school teachers. In Chapter 3, the dimensionality of measurement numeracy was investigated, to establish construct validity in measuring students’ measurement numeracy. Confirmatory factor analyses were performed on students’ test responses to ascertain whether estimating measures, understanding relationships within the metric system, calculating with scale and calculating length, area, and volume reflect one and the same student skill, or separate student skills. Chapter 4 describes the development of two lesson series: one with a pure deductive didactic approach, and one with a pure inductive didactic approach. In Chapter 5, an intervention study, in which video-taped lessons were used to measure classroom interaction time and teachers’ question types, is described. Results were used, in repeated measures ANOVA analyses, to estimate the effects of the didactic approach and the teacher on classroom interaction time and on the teachers’ question type. Finally, Chapter 6 describes the differential effects of the two didactic approaches on student performance. Students’ measurement numeracy was measured before and after the lesson series. Results were used to estimate the effects of the didactic approach and the teacher on students’ measurement numeracy.

7.1 Conclusions

7.1.1 Test construction and measurement of numeracy sub-skills
Confirmatory factor analyses (CFA), on student responses on a selection of items previously developed by the Rotterdam School of Education for a measurement numeracy course, were used to verify that items that were supposed to measure a specific sub-skill of measurement numeracy, indeed reflected only that sub-skill. In a pilot study, items with a high item-total correlation (to maximize reliability), and items varying in p-value (for differentiation purposes) were selected. This item selection was also used to measure the measurement numeracy of students in the quasi-experiment. The results of the CFA were that understanding
relationships within the metric system, calculating with scale, and calculating length, area, and volume did not reflect one and the same student skill, they reflected distinguishable student sub-skills. The hypothesized factor estimating measures could not be identified (although this might be due to small sample sizes and dichotomous scoring, see Chapter 3, half of the factor loadings of estimating measures items were non-significant and the average factor loading was below .4). If estimating measures were a separate student sub-skill, we would have found shared variance in the analyses of student responses, but we did not. A 3-factor model with correlated factors metric system, scale calculations, and length, area, and volume calculations fit well and better than a 1-factor model. This means that it may be beneficial to differentiate between these sub-skills in testing and in mathematics classes. Mathematics teachers might therefore consider paying attention to these distinct aspects in their tests (test each sub-skill separately) and courses. Naturally, students also need to be able to solve multifaceted problems (i.e. problems which require students to have two or more sub-skills). However, if tests and problems in classes consist only of these multifaceted problems, it becomes difficult for teachers to find the cause of low performance (which sub-skill has been insufficiently developed?).

7.1.2 Development of the lesson series

Two measurement numeracy lesson series were developed for prospective elementary school teachers: one with a pure deductive didactic approach and one with a pure inductive didactic approach. We started with a literature review and structured interviews with experts in the field of mathematics education. From there, PowerPoint blueprints and teacher manuals were developed in focus groups. Since measurement numeracy contains (at least) three distinguishable sub-skills, separate lessons were developed for each sub-skill. Lesson 2 was developed around understanding relationships within the metric system, lesson 3 was developed around calculating with scale, and lesson 4 was developed around calculating length, area, and volume. Lesson 1 was developed as introduction to the course, and also covered estimating sizes (this was not part of the research, but the teacher training curriculum required the students to be tested on this aspect, as well). Lesson 5 was developed as a summary and a formative assessment (i.e. discussion about practice exam items). The results are two lesson series on measurement numeracy, with PowerPoint sheets for every lesson and a teacher manual. Table 4.2 and Table 4.3 (Chapter 4) show the focus groups advice on what
teachers and students should be doing for each lesson (for each sub-skill) for the two didactic approaches.

The inductive idea was to choose a realistic context and problem as a playground for students to discover mathematical procedures and rules. The focus group for the inductive didactic approach suggested that, to discover general procedures, the teachers should guide student groups, in their discussions amongst themselves about the problem. After discussing several solutions, student groups should come up with general procedures to solve the problem. Teachers should ask students to share and explain their ideas, and ask other students to respond, so that students can individually select their preference for a procedure. Teachers who use an inductive didactic approach must have a solid teaching and mathematics background, as it is more complex to guide students’ reinventions and appropriately respond to their ideas than it is to explain procedures you already are familiar with. Students will not always be able to discover suitable procedures (Pierson, 2008) and students may ask for explanations (Gravemeijer, 1994) instead of engaging in each other’s mathematical ideas (Staples, 2007). To anticipate that, teachers were instructed to use scaffolding.

The deductive idea was to explain rules and to model procedures, step by step, before students individually apply them in different contexts. The focus group for the deductive didactic approach suggested that the teachers should ask students to share and explain their answers, and that the teacher should correct the explanations and answers until they are precise and clear. Teachers who use a deductive didactic approach must be able to explain how and why procedures work, to deepen students’ procedural knowledge, and to anticipate and address misconceptions (VanLehn, 1990). Students’ engagement in classroom interaction on how to use the teacher’s procedure (Van de Craats, 2007) may not always be optimal (for example, because applying the procedure is either too hard or too easy for some students). To anticipate that, teachers were instructed to use different kinds of questioning, depending on individual student’s level of understanding.

A limitation of this study is that only two experts were interviewed and that focus group members were all teacher educators of one teacher college. Other teacher educators or elementary school teachers might have shed another light on the development process. As the lesson series were developed for prospective elementary school teachers, who already learned about measurement earlier, and might therefore be more able to rediscover procedures than elementary school pupils, generalizability of the results (the lesson series) to elementary schools might be questionable.
Although this study resulted in lesson series with a pure deductive and one with a pure inductive didactic approach (which was needed to estimate the differential effect of the two approaches), in practice it is probably prudent for teachers to build their own personal mixture of these approaches, and to switch approaches depending on students’ needs. After the lesson series, teachers reported that students varied in approach preference, that discussions in the inductive didactic approach were time-consuming, and that many students who were taught inductively asked teachers to provide explanations instead of trying to rediscover procedures for themselves. Teachers also felt that, for some convenient procedures (with scale and the metric system), a deductive rather than an inductive didactic approach might be more suitable, given the apparently higher efficiency gains.

7.1.3 Effects on classroom interaction
The results in Chapter 5 show that an inductive didactic approach induced significantly more classroom interaction time (student talk) than a deductive didactic approach. When teachers explained a procedure that students had to follow in exercises, classroom interaction was less (40% of the time) than when teachers asked students to come up with their own procedures (52% of the time). Teachers also asked significantly more stimulating questions when they used an inductive didactic approach ($M=1.50$), compared to when they used a deductive didactic approach ($M=0.22$) (coding: 0 for controlling questions, 1 for equal, 2 for stimulating questions). Apparently, teachers were more inclined to induce students’ reflection and critical thinking in the inductive didactic approach than they were in the deductive didactic approach. These findings are interesting, since classroom interaction in mathematics classes allegedly has a positive effect on student outcomes in elementary schools (Slavin & Lake, 2008), and probably also on teacher college student outcomes (see Chapter 1). There was no teacher effect of the didactical approach on either classroom interaction time or the teacher question type. However, had teachers been free to make their own choices, and were not instructed to follow the teacher manual (which they apparently did very well), perhaps there might have been differences between teachers.

7.1.4 Effects on student performance
In Chapter 6, the effect of the didactic approach (inductive versus deductive) and the teacher on student performance was estimated. There was a significant effect of the course: learning gains from pretest to posttest were significant (in both approaches). The pretest and the
WISCAT score had a significant effect on the posttest score. However, ANCOVA analyses show that, after controlling for the pretest score and the WISCAT score, the didactic approach and the teacher did not have a significant effect on students’ measurement numeracy. This also means that these results cannot confirm or deny the idea that a deductive didactic approach might be a better choice for a group of (mathematically) low performing students (Milo, 2003; Timmermans, 2005).

Also, the type of questions the teacher asked, and the classroom interaction time did not have a significant effect. Although an inductive didactic approach induced more classroom interaction time and more stimulating teacher questions, students who were taught with the inductive didactic approach did not perform better than students who were taught with the deductive didactic approach. The teacher’s preference for a particular didactic approach, student behavior during class, previous education of the student and his parents, mathematical background, gender, age, and the student’s home language did not have a significant effect (although there was a small significant effect of the students’ mother’s education for the area aspect).

We expected to find an effect of classroom interaction on student performance (Slavin & Lake, 2008), but we did not find effects of teacher question type or classroom interaction time. Even though we statistically corrected for student attendance, for which we found no effect, perhaps the low student attendance (around 50%) influenced the relationship. The difference between classroom interaction time in the two didactic approaches may have been too small (40% of the time versus 52% of the time). Perhaps there is only an effect if the classroom interaction time in one of the conditions is much lower, for example below 20%, or much higher, for example above 80%.

7.2 Limitations
An initial limitation of this study is that all participants in the sample were students of one Rotterdam teacher training college. These students differ from the national population in ethnicity (more immigrants than the national average) and previous education (lower than national average). Furthermore, the WISCAT score in our sample was lower than the national average, and that variable was related to the posttest score. However, these three variables did not interact with the condition. The sample differs from elementary school children (age and previous education are the most obvious differences). Since students in the sample had already previously learned about measurement, chances are they were more likely to be able
to (re)discover procedures, and to come up with different kinds of procedures, than elementary school pupils. Therefore, generalization of the conclusions a) *an inductive approach induced more classroom interaction time*, and b) *student performance did not differ between an inductive and a deductive approach* to elementary school mathematics learning and teaching is questionable.

In this study we focused on a short lesson series. If the deductive didactic approach were to be applied over a longer period, teachers could choose to elaborate on certain aspects in order to deepen students’ procedural knowledge (Star, 2007). If the inductive didactic approach were to be applied over a longer period, teachers might not need to revert to scaffolding as quickly as suggested in this study. Teachers could then make room for student elaborations on other aspects of mathematics (or even other subjects) as well, because the time spent on these subjects or aspects would not be lost time.

We measured classroom interaction time and the teachers’ question type (Nelissen, 2002), but we did not measure quality of instruction (Hill, Umland, Litke, & Kapitula, 2012; Learning Mathematics for Teaching Project, 2011) or classroom discourse (Pierson, 2008). Although teacher effects on learning gains were not significant (see Chapter 6), as was a mismatch between the teacher’s preference and the didactic approach, and teachers followed instructions in both approaches just fine, quality of instruction and classroom discourse might have been different in the two approaches.

Random assignment of students to groups (to a teacher, or to an approach) was not possible, because students were in pre-existing groups that were pre-assigned to a teacher. However, student characteristics were reasonably equal across conditions (see Chapter 2). The design was unbalanced: not every group had the same number of students. However, overall, the difference in number of students in the two approaches was minor. Lesson attendance was rather low (lessons were not compulsory for students). The reasons for not attending lessons were not recorded (perhaps students were not motivated, perhaps they did not need the lessons). However, lesson attendance was controlled for in the analyses, and the effect on student performance was not significant.

The interrater reliability of the question type coding was rather low. However, the effect sizes, of the condition on the question type and on the classroom interaction, were extremely large. Even though an inductive didactic approach induced more classroom interaction time, and more stimulating questions, compared to a deductive didactic approach (see Chapter 5), no significant measurement numeracy improvement difference was found between the two approaches. However, the mean classroom interaction time in our
experiment was rather high: 48 per 120 seconds in the deductive didactic approach, and 62 per 120 seconds in the inductive didactic approach (see Chapter 5). Therefore, this study could not estimate the effect of very low classroom interaction intensity. One might think that differences in ability to teach with an inductive or deductive didactic approach matter, but a fidelity check showed that the teachers followed instructions just fine. We did not find a teacher effect, but this might be due to the instructions (didactic approach, teacher manual, PowerPoint sheets) they were given (all teachers followed the instructions for teaching quite well; although teacher 4 made some other choices, he complied reasonably with the instructions, see Chapter 5). If the aim was to estimate the teacher effect on measurement numeracy and on classroom interaction, it would have been necessary for teachers to have more freedom of choice in their classes.

Finally, classroom interaction was measured only on a group level. Perhaps measuring on student level would have given significant results (i.e. does student performance relate to his own active engagement in classroom discussions?).

7.3 Future directions
The findings in this dissertation, on the differential effects of an inductive and a deductive didactic approach on classroom interaction and on the absence of differential effects on student performance, raise several questions for future research. In the general introduction, the complexity of teacher education was described. The prospective teacher learns about pedagogy, about mathematics, about how to teach mathematics, and how to adapt to pupils’ educational needs. Teacher educators teach prospective teachers about numeracy, because a high level of numeracy is required in order to learn how to teach mathematics (Hill et al., 2005). The proof of the pudding, however, is not only in the test results, but most of all in the way prospective teachers teach mathematics in elementary schools (Hill, Umland, Litke, & Kapitula, 2012). Although the area of research into the complexity of teacher training college mathematics is wide open, this study focused only on enhancing prospective teachers’ measurement numeracy, by developing lesson series, and by estimating the effects of inductive and deductive teaching on student performance and on classroom interaction.

Results of this study show that an inductive didactic approach led to more classroom interaction and more stimulating teacher questions. Since improving classroom interaction allegedly has a positive impact on student performance (Slavin & Lake, 2008), an inductive approach might be advisable. However, classroom interaction can probably also be improved in a deductive approach. In a deductive approach, for example, teachers can be instructed to
use as many stimulating questions (versus check questions) as possible. Regardless of the didactical approach, it would be valuable to search for possibilities to improve classroom interaction, in such a way that student performance is optimized. Student performance can be interpreted in various ways, such as in different kinds of knowledge.

The distinct difference between conceptual and procedural knowledge is easy to comprehend (through examples of being able to follow a procedure – with or without misconceptions – without knowing what you are doing), yet hard to describe (Hiebert, 2013). Conceptual knowledge requires relationships between pieces of information in the learner’s network of knowledge. Procedural knowledge requires step by step instructions on how to complete a task. A learner with only procedural knowledge is dependent on how problems are formulated: if he does not recognize the problem, he will not know that he is supposed to apply the procedure. For example: if a learner has learned to subtract in problems like \( 12 - 5 = ? \), he will not automatically be able to solve problems like \( 12 - ? = 5 \), or \( 5 + ? = 12 \). In the context of measurement numeracy: suppose a student has learned procedures for converting units, and for calculating volume. Then we can assume that he can solve problems like \( 23 \text{ dm} = \ldots \text{ m} \), and calculate the volume of a swimming pool with \( \text{ length} = 25 \text{ m}, \text{ width} = 12 \text{ m}, \text{ height} = 2\text{ m} \). However, he will not automatically be able to solve problems like: The volume of a pack of lemonade is 1.5 liters. The pack is 7.5 cm long, and 1 dm wide. Calculate the height of the pack. The last problem seems to test conceptual knowledge, but what if the teacher had practiced a procedure for these kinds of problems? (i.e. step 1 is to convert length and width measures to dm, step 2 is to divide the volume by the product of the length and width). What happens in the classroom seems to determine whether conceptual or procedural knowledge is tested. Furthermore, the difference between the two types of knowledge fades even more in the terms ‘procedural understanding’, ‘conceptual understanding of procedures’, and ‘mindful execution of procedures’ (Star, 2005). There is no clear correspondence between these different types of knowledge and the two didactic approaches in this study. After all, in the deductive approach, students solve problems after the teacher has explained the concept and the procedure; in the inductive approach the students solve problems, and then rediscover concepts and procedures. Both types of knowledge are part of both didactic approaches. The difference is the order in which the two types of knowledge come into play. Since literature about the best order (Fyfe, DeCaro, & Rittle-Johnson, 2014), and about appreciation for procedural knowledge (Star, 2005) is inconclusive, further research is needed.

Colleges and specific courses may differ in learning objectives. If an objective is: learning how to perform calculations of area (where the teacher knows example items), a
teacher might choose to have students practice procedures. If an objective is: learning about the concept of area (where the teacher knows no example items), a teacher might choose another approach. Since student results often play a role in performance evaluations of teachers and colleges, teaching procedural knowledge (teaching to the test) might be more attractive than teaching conceptual knowledge. It looks like a catch-22 situation: on the one hand, conceptual knowledge is desirable and national tests might prove to secure the overall knowledge level, and on the other hand, teachers and colleges are evaluated by student scores on tests that can be done by using procedural knowledge, if teachers teach procedural knowledge in various contexts (e.g. introducing national performance tests may be counterproductive, Van Zanten & Van den Brom-Snijders, 2007). Research on how to avoid this catch-22 may be valuable.

It would also be valuable to find ways to measure the effects of different approaches on other aspects: not only on a standard mathematics test, but also on other skills like transferring to other contexts or other mathematical domains, or explaining why one strategy is better than another. Since the inductive lesson series contains possibilities for students to discover procedures, students will not only learn about these procedures, but also about unstructured problem-solving and how to respond to other students’ ideas. Research into these effects (discovering what exactly separates deductively taught students from those who were taught inductively) may be of value, too. For example: which types of problems are expected to be solved more easily by the first or second group of students?

Classroom interaction time changed significantly over lessons: there was more interaction in the lesson on scale calculations (54%) than in the other two (metric system: 38%, and area calculations: 46%). Perhaps mathematical subjects differ with regard to the classroom interaction time they induce. The relationship between the mathematical subject and classroom interaction time may also be influenced by how much students already know about the subject. Further research is needed to specify the relationship between the mathematical subject and classroom interaction time. Furthermore, the contents of the classroom interaction (how do the students contribute to the classroom interaction, how does the teacher engage in student ideas?) were not under investigation in this study. As teachers ask more stimulating questions when they use an inductive approach, and deeper discussions (and the way teachers stimulate these deep discussions) on mathematically rich contexts might influence student performance (Hill, Umland, Litke, & Kapitula, 2012; Pierson, 2008; Star, 2005), it may be valuable to estimate the differential effect of the content of classroom
interaction (i.e. the mathematical richness of student participation, and the way teachers respond) on student performance between an inductive and a deductive approach.

Students talked more about mathematics in the inductive approach, than in the deductive approach. However, in group work, chances are that some students tend to talk more than others. In this study, classroom interaction was only measured at group level. Perhaps measuring at student level would produce different results. Is a student’s performance related to his active engagement in classroom discussions? If that relationship is positive, it might be prudent to put extra effort into stimulating every single student to actively engage. Teachers can achieve this by using, for example, some aspects of cooperative learning, in which students have specific roles: two students share ideas, one students ask questions to deepen the ideas, and one student reports on the discussion in a whole class evaluation, next time the roles change, so that every student can actively practice every role.

In the Rotterdam School of Education, teacher educators have less knowledge about their students’ mathematical abilities than elementary school teachers have about their pupils, because, unlike in elementary school, teacher educators see their students only 50 minutes per week. Feedback on homework assignments is not provided, because the costs are considered to be too high (it is time consuming). Therefore, there is hardly any differentiation between students in numeracy classes; suggestions for leading productive mathematics discussion, like anticipating, monitoring, selecting, sequencing, connecting (Smith & Stein, 2011) can hardly be followed. Perhaps ways can be found to lower the costs (only give feedback on key parts of the assignment, only select student work for the class, …). Research on the effect of these suggestions for leading productive mathematics discussion on prospective teachers’ measurement numeracy may be valuable.

Finally, further research is needed on what estimating measures is about and how it correlates with the sub-skills of measurement numeracy. Is estimating measures one skill, or a subset of sub-skills (for example estimating length, estimating area, estimating volume, and estimating weight)? It might not be safe to assume that students estimate different measures using one and the same strategy (i.e. think of a reference measure and compare it with the object that requires estimation), and furthermore, Morewedge and Kahneman’s (2010) research showed that estimations are biased because of priming and personal environments. How do alleged estimating measures sub-skills relate to the three sub-skills that were identified in this study? If some estimating measures skill relates to another sub-skill of measurement numeracy, it might be effective to teach those sub-skills together. If, however, estimating measures is not related to any other numeracy skill, policy makers might
reconsider the place of estimating measures in the curriculum. Perhaps it is primarily helpful in everyday life, and desirable to check if answers are within a logical range, as is stated in the Dutch knowledge base for mathematics (Van Zanten et al., 2009), but not necessary for developing other measurement numeracy skills. However, following Freudenthal’s ideas, Dutch textbooks for measurement didactics focus on the sense of measures; they view measurement as a way to estimate sizes, and place extra emphasis on the ability to imagine units and measures in different contexts (De Moor, 2005).

### 7.4 Final remarks

In the inductive didactic approach, teacher questions were more stimulating and there was more classroom talk, and teachers also reported that students were more engaged. However, student performance did not differ between didactic approaches. What will make a difference? More differentiation in approach, by tuning in on student needs? Using student work for classroom discussions? Motivating students to work harder on assignments, to be more actively engaged during classes? We conclude that further research is needed on the effect of distinct aspects of classroom interaction on numeracy improvement, in order to empirically substantiate claims of positive effects on learning gains.