Skrifter från SMDF, Nr. 18

Mediating mathematics

Proceedings of MADIF 14
The fourteenth research conference of
the Swedish Society for Research
in Mathematics Education
Örebro, March 19–20, 2024

Editors:
Johan Häggström, Cecilia Kilhamn, Linda Mattsson,
Hanna Palmér, Miguel Perez, Kerstin Pettersson,
Ann-Sofi Röj-Lindberg, Anna Teledahl

SMDF Svensk Förening för MatematikDidaktisk Forskning
Swedish Society for Research in Mathematics Education

Skrifter från Svensk Förening för MatematikDidaktisk Forskning, Nr 18

ISBN 978-91-984024-7-6
ISSN 1651-3274

SMDF
Svensk Förening för Matematikdidaktisk Forskning
c/o Nationellt Centrum för Matematikutbildning
Göteborgs universitet
Box 160
SE 40530 Göteborg

Swedish Society for Research in Mathematics Education

For information see web page http://matematikdidaktik.org

Published electronically only

Preface

The 19th - 20th of March 2024, the Swedish Society for Research in Mathematics Education (SMDF) convened 107 scholars at the Örebro University for the fourteenth Swedish Mathematics Education Research Conference – MADIF 14. The conference, themed *Mediating mathematics*, marked a significant milestone: for the first time in its history the scholarly program commenced on the opening day. Professor Nathalie Sinclair, Simon Fraser University, Canada, initiated the conference with a captivating plenary lecture titled *Towards embodied validity in mathematics education research*. Her thought-provoking presentation ignited fresh perspectives on learning mathematics and sparked engaging discussions during the subsequent conference dinner. On the following day professor Elham Kazemi, University of Washington, USA, welcomed participants into the mathematics classroom through her inspiring plenary lecture on *Teachers learning together with and from students*. Her excellent presentation set the stage for the remainder of the conference. The program committee extends heartfelt gratitude to the invited speakers for framing the conference with their insightful and inspiring lectures.

Acknowledgments

The program committee would also like to thank all participants of MADIF 14 for fostering an environment where research perspectives and thoughts were exchanged in an open and positive manner. Your contributions have made this conference a shining example of scholarly collaboration and intellectual growth.

Furthermore, we would like to thank all of you who submitted papers, short communications, workshop contributions, or symposia to MADIF 14. Your valuable contributions enriched the conference by fostering engaging interactions and inspiring the writings presented in this volume. The eleven papers are research reports presenting original work that has not been previously published. Thirty-one abstracts of short communications summarize ongoing research projects or specific aspects of work presented during the conference. They provide glimpses into ongoing investigations. Three workshops focus on research topics or questions that are still under development. They serve as collaborative spaces for exploring emerging ideas. The symposium proposal consists of documents presenting two or three studies centered around

a common theme. These studies were presented and debated during symposia sessions. This volume constitutes the proceeding of this conference, providing a glance at current exciting research in mathematics education in Sweden and other Nordic countries.

At MADIF 14, all presentations were followed by discussions led by allocated discussants. The program committee would like to recognize the fine work done by these scholars:

> Andreas Bergwall, Andreas Eckert, Angelika Kullberg, Anna Teledahl, Camilla Björklund, Cecilia Kilhamn, Daniel Brehmer, Hanna Palmér, Helena Roos, Iben Maj Christiansen, Jan Olsson, Jorryt van Bommel, Julia Tsygan, Jöran Petersson, Kerstin Larsson, Kristofer Sidenvall, Laura Fainsilber, Linda Marie Ahl, Lotta Wedman, Maria Lundkvist, Ola Helenius, Olov Viirman, Robert Gunnarsson, Tomas Bergqvist, Tuula Koljonen, Ulrika Ryan and Ulrika Wikström Hultdin.

Moreover, we are very thankful for the great work done by our reviewers. Some reviewers presented papers at the conference, others contributed to our community as external reviewers. The external reviewers were:

> Andreas Ebbelind, Andreas Ryve, Andreas Tamborg, Anna Pansell, Anne Tossavainen, Björn Palmberg, Elena Nardi, Eva Norén, Frode Rønning, Johanna Pejlare, Jorryt van Bommel, Jöran Petersson, Lars Madej, Lovisa Sumpter, Magnus Österholm, Martin Carlsen, Niclas Larson, Olov Viirman, Paula Valero, Per Nilsson, Peter Nyström, Samuel Sollerman, Yukiko Asami-Johansson and Yvonne Liljekvist.

The program committee would also like to express their gratitude to the organisers of Matematikbiennalen 2024 for their support in making this conference possible. Additionally, we appreciate the Örebro University for generously providing the necessary premises and infrastructure. Your collective efforts significantly contribute to disseminating research and exchanging scholarly ideas in mathematics education in Sweden.

Finally, as the chair of the program committee for MADIF 14, I, Linda Mattsson (Blekinge Institute of Technology), would like to thank the excellent members of the committee, Johan Häggström (University of Gothenburg), Cecilia Kilhamn (University of Gothenburg), Hanna Palmér (Linneaus University), Miguel Perez (Linneaus University), Kerstin Pettersson (Stockholm University), Ann-Sofi Röj-Lindberg (Åbo Akademi University) and Anna Teledahl (Örebro University). The synergy of diverse competencies within this committee has been remarkable. Each member's unique contributions have woven together a tapestry of excellence. As the committee chair, I witnessed everyone go above and beyond and surpass any expectations.

Review process of and reference to proceeding

In a rigorous two-step review process for presentation and publication, all papers were peer-reviewed by three researchers. Contributions viewed as short communications or symposium proposals were reviewed by members of the programme committee. The MADIF Proceedings are published by SMDF (skriftserie från Svensk Förening för MatematikDidaktisk Forskning), which since 2010 has been designated scientific level 1 in the Norwegian list of authorised publication channels available at Norwegian register for Scientific Journals, Series and Publishers (kanalregister.hkdir.no).

Due to climate considerations SMDF has decided to publish only a digital version of this proceeding. When referring to this volume please write the following.

Häggström, J., Kilhamn, C., Mattsson, L., Palmér, H., Perez, M., Pettersson. K., Röj-Lindberg, A.-S. & Teledahl, A. (Eds.) (2024). *Mediating mathematics. Proceedings of MADIF14*. SMDF.

The MADIF conference proceedings are published in the SMDF writing series. From MADIF 9 they have also been published electronically and are available on the SMDF website (matematikdidaktik.org).

The next conference, MADIF 15, will be held at the University of Gothenburg in 2026. Whether you are a novice explorer in this field or a seasoned attendee of past conferences, we eagerly anticipate your presence in Gothenburg. Let's continue our scholarly journey together!

MADIF 1, 1999, Stockholm
MADIF 2, 2000, Göteborg
MADIF 3, 2002, Norrköping
MADIF 4, 2004, Malmö
MADIF 5, 2006, Malmö
MADIF 6, 2008, Stockholm
MADIF 7, 2010, Stockholm
MADIF 8, 2012, Umeå
MADIF 9, 2014, Umeå
MADIF 10, 2016, Karlstad
MADIF 11, 2018, Karlstad
MADIF 12, 2020, Växjö
MADIF 13, 2022, Växjö
MADIF 14, 2024, Örebro

The program committee for MADIF 14
by chair Linda Mattsson

Contents

Developing a framework for assessing students' mathematical writing

Cecilia Kilhamn and Anna Teledahl

Assessing students' writing in mathematics is complex, and teachers are often left with limited support to identify and categorise quality in texts that report on problem-solving. This study is part of a larger project in which teachers and students work on discussing and improving the quality of students' mathematical writing. For this paper, we have developed a framework for assessing the quality of students' texts by operationalising three aspects of quality. One hundred and fifty-four student texts were analysed to develop definitions of quality levels and validate the framework. Findings show that clarity, efficiency, and use of mathematical notation in student texts can be scored independently for the problem description in students' problem-solving texts, that differences in quality can be discerned, and that a progression in quality over time becomes visible. Some challenges are discussed.

Assessing the quality of students' written mathematical work is difficult for mathematics teachers. There are several reasons for this (Casa et al., 2016; Powell et al., 2017). Firstly, it is worth noticing that assessment of students' written documentation of problem-solving is seldom focused on the quality of their writing, separate from the quality or effectiveness of their (choice of) problem-solving strategies, making it difficult to provide students with feedback focusing exclusively on their writing skills. Secondly, there are very few examples of support for teachers in the form of research findings or curricular standards that describe aspects of mathematical writing in ways that can be used to assess different levels of quality in students' written work. During the last two decades, research in mathematics education has paid much attention to communication, primarily highlighting oral communication, i.e., organising opportunities for students to communicate with, in, and about mathematics in group collaborations and/or whole class discussions. Research on students' written communication appears to have taken a back seat.

In a project focused on students' written communication, we set out to investigate students' writing by developing a teaching design that allows teachers and students to analyse and discuss issues of quality (Teledahl et al., 2023). From

Cecilia Kilhamn, NCM, Gothenburg University
Anna Teledahl, Örebro University and Dalarna University

these discussions and informed by previous research, we have constructed a framework in which we suggest four elements that should feature in students' written documentation of problem solving and three aspects of quality for these elements, visible in figure 1 and described in more detail in following sections.

Elements / Quality	Problem description	Calculations	Arguments, justifications	Conclusion
Efficiency				
Clarity				
Notation				

Figure 1. *A framework for discussing quality in student written documentation*

This paper reports on the operationalisation of the framework as a tool for assessing the first element across the three aspects of quality. In a sense, it is about defining quality in student-written texts to develop a framework that researchers and teachers can use to assess quality and see progression. Our research question is thus: How can quality in the three aspects of the framework be described so that they can be used to assess problem descriptions in grade 5 student texts?

Background

We suggested initially that assessing the quality of students' mathematical writing is complex and that one of the reasons was that there is little support from standards or research literature. Two of the most important reasons for this are that a) writing in mathematics combines several different communicative resources to create meaning (O'Halloran, 2005), and b) what is considered good writing depends both on the local context and on conventions adopted by the global community (Barwell, 2018).

Since conventions regarding writing in mathematics are mainly concerned with mathematical notation, research has often focused almost exclusively on students' use of mathematical symbols (Hughes et al., 2019; Teledahl et al., Forthcoming). In school mathematics, and especially in written reports of problem solving, students need to create a narrative that lets a reader understand what the problem is, what the premises are, what calculations and other procedures are carried out and what the arguments for these are as well as what the conclusion is. Students cannot create such a narrative using only mathematical notation. Instead, as O'Halloren and others have described, they are required to combine communicative resources such as natural language and visual representations with mathematical notation in ways that clearly and efficiently convey necessary information and create a comprehensive narrative (O'Halloran, 2005; Namkung et al., 2020; Powell et al., 2017; Steenrod et al., 1973).

The framework

The framework (figure 1) is informed by previous research findings on writing in mathematics (see, for example, King et al., 2016; Kosko & Zimmerman, 2019) together with research on communicative competence in general (Rickheit et al., 2008) and several pilot studies in the initial phase of the project. The original descriptions of the four elements are as follows:

1. The *problem description* includes a description of the conditions consisting of two parts: i) premises and facts, i.e., the mathematically relevant information given in the problem and assumptions that need to be made, and ii) what is asked for in the problem, the question to be answered.
2. *Calculations* include all mathematical procedures needed to solve the problem, such as structuring, representing, modelling, calculating, or solving equations.
3. *Arguments and justifications* include making explicit the reasons behind the different steps in the solution, describing why specific calculations are made and where the included numbers and variables come from, as well as the logical steps that lead to a conclusion, applying logical reasoning and using implications.
4. A *conclusion* is the final response that relates to the original problem.

We expected each of these elements to appear in qualitatively different forms. Based on a literature review (Teledahl et al., Forthcoming), we initially identified four aspects of quality: clarity, efficiency, use of mathematical notation, and appropriateness. Clarity is related to readability and comprehension. *Efficiency* is described as communication that achieves its goal with as little effort as possible. *Mathematical notation* is often seen as the ultimate example of efficiency, but to be understood, it needs to be well-chosen and correctly used following traditional conventions. The element *appropriateness*, defined as the choice of linguistic or other resources best suited to the context and the expected reader, was later removed and is here incorporated in each of the three remaining aspects (Teledahl, 2023; Teledahl et al., 2023). Appropriateness alludes to ”speaking the language of the classroom” and making the different elements of the text understandable. Thus, it overlaps with the ideas of both clarity and efficiency. It is also a feature of mathematical notation, as we expect students to use mathematical notation appropriate for their grade level. Consequently, in this paper, we have dealt with issues of appropriateness in terms of what expectations we have at the specific grade level in relation to the first three aspects rather than analysing it separately as one of four aspects of quality.

The teaching design

The data for this study is collected from a research project in which we developed a teaching design that would ensure a separation between the two processes of (i) solving a mathematical problem and (ii) communicating a solution to the problem in writing. Such a separation creates an opportunity for teachers and students to have discussions that focus exclusively on the quality of students' writing.

The first part of our design is a problem-solving lesson, enacted by each teacher in a way the students are used to, but ending with a whole-class session in which students present and discuss different solution strategies so that, in the end, most of the students will have grasped the solution and at least one strategy for solving the problem. They are then required to produce written documentation of their problem-solving, but instead of asking students to ”describe their thinking”, teachers tell them to ”describe and justify any one of the ways the problem can be solved so that someone who has not solved the problem can understand the solution”. These documentations, henceforth called *student texts*, are collected by the teacher.

The second part of the design, in most cases as a second lesson separated in time from the first, is aimed at discussing the quality of the student texts, i.e. the focus is exclusively on the communicational quality, not the choice of problem-solving strategy. During this lesson, teachers and students analyse some of the students' texts and discuss their merits. In the organisation of the discussions, the teachers have access to the framework described in figure 1 (initially with appropriateness added as a fourth aspect as related above). Teachers can choose to emphasise different aspects during the discussion depending on the texts produced in each classroom. After the quality discussion, students produce another text documenting the same problem, hopefully a better version than their first try.

Method

Data collected for this study consist of student texts from five different grade 5 classes in a Swedish municipality (classroom A, B, C, D, E). As participants in the project, the teachers and students went through the teaching design working with different problems six or seven times during one semester. Each teacher collected student texts at two points in time for each problem – at the end of the problem-solving lesson and again after the quality discussion. Mostly the texts were produced as group or pair work. In addition, the teacher in classroom E let her students solve the first problem again at the end of term, collecting individual student texts at a third data point for that problem. In total, we analysed 135 student texts on the first problem called the Animal problem: 41 are from the first data point, 72 from the second, and 22 from the third data point in classroom E. In addition, the validation phase includes another 19 student texts on two other problems from three different classes.

The Animal Problem
In a village there are four different kinds of animals: pigs, sheep, hens, and cows. Every fourth animal is a pig. One out of eight animals is a sheep. Half of all the animals are hens. The rest, 50 animals, are cows.
How many animals of each kind are there in the village?

Method of analysis

To answer our research question, the framework had to be operationalised so that different levels of qualities of each aspect could be identified. We started to work with the first element, *problem description*, with the initial idea that generic qualities could be described and later used for the other elements. This initial work is reported here.

In a previous study, Teledahl (2023) used multimodal discourse analysis (Jewitt, 2011) to analyse written documentations from grade 6 and grade 11. She looked specifically at how different semiotic resources were used to create clarity and efficiency in these texts. Based on her initial findings, the work of operationalising the framework was conducted through a bottom-up approach, taking one of the three quality aspects at a time. Qualities that appeared as similar in different student texts were thematically collected into categories in an iterative process as described by Gläser-Zikuda et al. (2020). We also made note of the communicative resources used in the written documentations, distinguishing between *linguistic resources*, such as complete sentences or linking words; *visual resources*, such as headings, boundaries, arrows, bullet points, or sequencing; *mathematical symbols*; and *graphical resources* such as charts, tables, or models. Since the goal was to be able to use the framework to assess quality and progression, the categories within each quality aspect were hierarchically organised .

Once initial categories had been identified, the whole data set for the Animal problem was assessed using these categories as levels of quality. When problems appeared, adjustments to the category descriptions were made to avoid overlap and ambiguity, after which the whole data set was re-assessed. One issue discussed was to what extent the categories were to be generically described versus specific to the problem or the element. For the framework to be useful, we expected some progression to be visible. After three iterations, final levels of quality were formulated for the element *problem description*, and assessments of the student texts were compiled. Finally, a validation of the resulting levels of quality followed, where 19 students' texts documenting other problem solutions were assessed.

Results

In this section, we first report how the final levels of quality came to be formulated for each for the three quality aspects. We then report the results from the assessment of all the texts using the framework, and separately for classroom E to show long-term progression. Finally, we describe the results of the validation phase.

Quality in problem description

Analysis of the problem descriptions in the students' written documentations of the Animal problem resulted in four levels of quality for each of the three aspects *clarity, efficiency* and *use of mathematical notation*. The levels are numbered from 0 to 3 indicating an increased quality for higher values. In addition, the choice of communicative resources was documented.

Some interpretations related to appropriateness for the specific age and problem were necessary. We expect students in grade 5 to clarify the different proportions of pigs, hen, and sheep and the fact that there are 50 cows. Differentiating between proportion and quantity is seen as essential. To be solvable, the problem description ideally also includes the fact that there are *only* four different kinds of animals, or that the 50 cows constitute *the rest* of the animals, but we do not expect grade 5 students to discern this detail as essential to declare in the description.

Levels of quality for Clarity

0 No introduction of the problem (premises can be implicit in the graph or calculations).

1 Partially correct and comprehendible conditions given, or all premises given but no question, or ambiguous descriptions.

2 Conditions are given in various places of the text or explicitly integrated with calculations. All essential conditions are present.

3 Separation, with all essential conditions given before the calculations.

Nearly all student texts include some kind of graphical representation of the Animal problem and sometimes it is difficult to separate the problem description from the solution process. Making a graphic representation is generally interpreted as the first calculation in the solution process. But if there is no other problem description, it is assessed as the problem description with clarity level 2 if the facts are clearly and explicitly stated as premises in the graph and as level 0 if they are not.

Levels of quality for Efficiency

0 Difficult to make out what the conditions are (not enough information to evaluate efficiency, e.g. missing, messy or ambiguous).

1 Description of conditions including redundant text or information, for example, copying the whole problem verbatim as given.

2 Efficient presentation of essential conditions using mainly natural language and visual resources.

3 Efficient presentation of essential conditions using appropriate mathematical, symbolic, or graphical resources.

It is worth noticing that a text can be assessed as efficient on level 3 even if the mathematical notation is not entirely correct. If, for example, the equal sign is misused but the intention is understandable at the grade level in question, it is assessed as efficient. Likewise, correct spelling and grammar are not essential for assessment on level 2.

Levels of quality for Mathematical notation

0 No mathematical notation is used except what was given in the original problem formulation.

1 Incorrect use of mathematical notation, or use of a graphical representation that does not represent the problem correctly. (The most common misuse found in our pilot studies was an incorrect use of the equal sign.)

2 Irrelevant or redundant mathematical notation used essentially correct.

3 Relevant mathematical notation used correctly.

Many students use fractions to represent the number facts given in the problem. Here, the difference between proportions and quantities comes into play, as well as the use of the equal sign. While still accepting some mathematical notation as appropriate to grade 5 students, some nuances are assessed as incorrect. The symbol 1/4 can be read as ”one fourth” or ”one out of four”. Ideally, it should be stated that it is ”one fourth of the total number of animals”. In most cases this is accepted as taken for granted unless there is cause for misunderstanding. Some examples of mathematical notation assessed as correct are: ”1/4 is pigs and 50 units are cows” (In Sweden, the word ”stycken” is used to specify discrete units of something); ”1/4 is pigs and 50 are cows”; ”Pigs: 1/4 … The rest, Cows: 50”. Notations assessed as incorrect are, for example: ”1/4 = pig, 50 = cows” (misuse of the equal sign) or ”pig: 1/4 … cows: 50” (ambiguous since the difference between proportion and quantity is not made clear).

Assessment of student texts

Results of the analysis of students' problem descriptions are presented first for the 72 + 41 texts describing solutions to the Animal problem from the participating five classrooms (figure 2), then separately for all the texts from classroom E to show progression (figure 3).

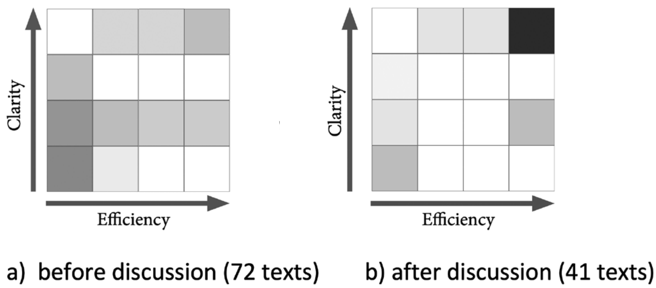

Figure 2. *Distribution of texts according to assessed levels of quality across the two aspects,* clarity *and* efficiency *for the Animal problem*
Note. The darker the colour, the higher the frequency

The results indicate that the framework was useful for assessing student mathematical writing quality. We can see that the students did embrace more efficient ways of communicating as a result of the teaching design focusing on quality aspects, and that their texts improved in terms of clarity. In the first diagram, many texts, are found on the lower left side on efficiency level 0 and clarity level 0 or 1. In the second diagram, most texts are assessed on level 3 for both aspects.

Assessing progression

We took a closer look at classroom E, where the students were given an opportunity to solve the Animal problem and report their solution in writing at the end of term. In total, 41 texts came from classroom E: seven after the first lesson (group work), twelve after the quality discussion (pair work), and twenty-two individual texts at the end of term. Figure 3 shows the distribution of texts across levels of quality for the two aspects, clarity and efficiency. In the seven texts collected after the first problem-solving lesson (figure 3a), mainly linguistic resources are used with few instances of mathematical notation. Only one text clarifies all the given conditions before the calculations and conclusion. After the quality discussion (figure 3b), almost every text clarifies the conditions using visual resources and mathematical notation, raising the level of quality for efficiency to level 3. Eleven out of twelve texts communicate efficiently using mathematical notation, albeit not always using the notation correctly.

After working with another six problems using the same teaching design, the students were again given the Animal problem to solve and 22 individually written documentations were collected (figure 3c). The predominant resources

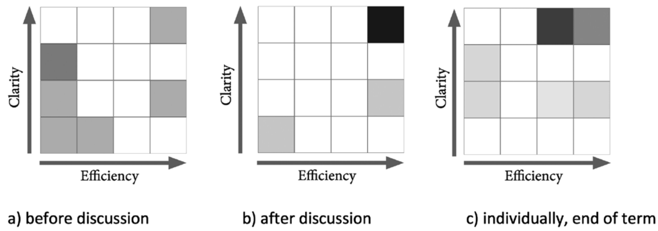

Figure 3. *Distribution of texts from classroom E according to assessed levels of quality across the two aspects clarity and efficiency for the Animal problem*
Note. The darker the colour, the higher the frequency.

in these texts are linguistic and visual (efficiency level 2). Only seven texts include appropriate mathematical resources, suggesting that over time, awareness of clarity has increased but that all students, individually, are still not fluent with the appropriate mathematical symbols and graphical resources.

Validation

To validate the operationalisation of the framework, problem descriptions of two other problems in 19 student texts were assessed by both researchers separately. A comparison showed total agreement in the assessment of quality when using the framework. However, two reflections were made, both related to mathematical notation.

In one problem, all students assigned letters to variables, for example, letting g and s represent the weight of gold and silver medals to generate the equation $g+g+s=340$ gram. In most cases, there was no explicit assignment of these letters, and thus it was interpreted as a case of incorrect use of mathematical notation assessed as level 1, while the efficiency aspect was assessed as level 3 since appropriate mathematical symbolic resources were used efficiently. Clarity was also assessed as level 3 with the argument that the essential premises concerned the relationships between the variables, not necessarily what they represented. However, one student introduced an ambiguity in the description by writing $g+g+s$ *is 340*g, with two different meanings for the letter g in the equation, thus assessed as clarity level 1.

In another problem, no mathematical notation except for the numbers given in the original problem were deemed appropriate to describe the problem clearly and efficiently, therefore assessed as level 0 for use of mathematical notation. This result implies that efficiency level 3 cannot always be reached for the problem description.

Discussion

Results from our project have shown that the described teaching design does, in fact, offer opportunities to focus exclusively on, and thereby develop, students' ability to communicate in writing. We have also seen that access to the framework offers opportunities to disassemble writing and discuss different aspects of quality, something we believe has been missing from teachers' practices.

The findings reported here seem to confirm that all three aspects of quality: *clarity, efficiency* and *use of mathematical notation*, can be operationalised to enable assessment of students' written documentations, at least concerning the problem description element. As suggested by Steenrod et al. (1973), a combination of communicative resources was valued as a quality of efficiency, and increasingly efficient when appropriate mathematical notation was included. In line with previous research, we could also see that incorrect use of mathematical notation could cause ambiguity, implying a need to highlight both how and when to use new mathematical symbols and representations. Furthermore, using the described levels of quality to assess student texts made it possible to identify differences and see progression over time. One implication of the validation phase is that all levels of all three aspects of the framework are not always applicable to every element. Correct use of mathematical notation is probably, in many cases, a more prominent aspect of the element *calculations and arguments* than of the problem description.

We have in this paper only developed the framework for the assessment of the first element, the problem description, enabling only an analytic view of that particulate element. When the aim is to create a comprehensive narrative, as Powell et al. (2017) suggested, the problem description needs to be separated from the solution, which is why a separation is valued higher. However, further research is needed to operationalise the framework for all elements to make a more holistic assessment possible.

Defining the levels described here turned out to be more challenging than we first expected, partly because the aspects efficiency and clarity interact with each other. One intriguing result was that we did not find any text assessed for clarity level 2 that was more than efficiency level 0. However, we did assess several texts on clarity level 1 as efficiency level 1, 2 or even 3. This can be explained by the fact that clarity level 1 is defined as being *partial*, and the part that is there could well be efficiently communicated. While premises integrated with the calculations might be efficiently communicated from a holistic point of view, it is rarely the case when assessing only the problem description.

The levels of quality described for the first element were not fully generic. However, they could serve as a useful starting point when applied to other elements, which will be the next step in our research to develop the framework further.

Acknowledgements

This paper is part of grant 2020-00066 under the Swedish Institute for Educational Research.

References

Barwell, R. (2018). Writing in mathematics classrooms. In A. Bailey, C. Maher & L. Wilkinson (Eds.), *Language, literacy, and learning in the STEM disciplines* (pp. 1001–1114). Routledge.

Casa, T. M., Firmender, J. M., Cahill, J., Cardetti, F., Choppin, J. M. et al. (2016). *Types of and purposes for elementary mathematical writing: task force recommendations*. mathwriting.education.uconn.edu

Gläser-Zikuda, M., Hagenauer, G. & Stephan, M. (2020). The potential of qualitative content analysis for empirical educational research. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 21 (1).

Hughes, E. M., Markelz, A. M. & Cozad, L. E. (2019). Evaluating various undergraduate perspectives of elementary-level mathematical writing. *International Journal of Science and Mathematics Education*, 17 (5), 1031–1048.

Jewitt, C. (2011). Different approaches to multimodality. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (pp. 28–39). Routledge.

King, B., Raposo, D. & Gimenez, M. (2016). Promoting student buy-in: using writing to develop mathematical understanding. *Georgia Educational Researcher*, 13 (2), 33–52.

Kosko, K. W. & Zimmerman, B. S. (2019). Emergence of argument in children's mathematical writing. *Journal of Early Childhood Literacy*, 19 (1), 82–106.

Namkung, J. M., Hebert, M., Powell, S. R., Hoins, M., Bricko, N. & Torchia, M. (2020). Comparing and validating four methods for scoring mathematics writing. *Reading & Writing Quarterly*, 36 (2), 157–175.

O'Halloran, K. L. (2005). *Mathematical discourse: language, symbolism and visual images*. Continuum.

Powell, S. R., Hebert, M. A., Cohen, J. A., Casa, T. M. & Firmender, J. M. (2017). A synthesis of mathematics writing: assessments, interventions, and surveys. *Journal of Writing Research*, 8 (3), 493–530.

Rickheit, G., Strohner, H. & Vorwerg, C. (2008). The concept of communicative competence. In G. Rickheit & H. Strohner (Eds.), *Handbook of communication competence* (pp. 15–62). Mouton de Gruyter.

Steenrod, N. E., Halmos, P. R., Schiffer, M. M. & Dieudonné, J. A. (1973). *How to write mathematics*. The American Mathematical Society.

Teledahl, A. (2023). Students' strategies for presenting the problem. In P. Drijvers, C. Csapodi, H. Palmér, K. Gosztonyi & E. Kónya (Eds.), *Proceedings of CERME13* (pp. 1695–1702). Alfréd Rényi Institute of Mathematics and ERME.

Teledahl, A., Kilhamn, C., Ahl, L. & Helenius, O. (Forthcoming). *Defining and measuring quality in students' mathematical writing: a systematic literature review.*

Teledahl, A., Kilhamn, C., Helenius, O. & Ahl, L. (2023). Writing down the solution separate from solving the problem. *For the Learning of Mathematics*, 43 (2), 41–46.

Mediating mathematics out of the lesson: Tamra's personal breaching experiment

Iben Maj Christiansen

Research on how teachers' instructional choices are framed by professional obligations provides insight into the processes of decision-making. This is particularly pertinent when a teacher sets out to change her teaching. An understanding of these processes can inform support for teachers as well as teacher education. This paper reports on a lesson in which Tamra – for the first time – is trying out a more explorative approach with her grade 7 learners. Tensions prevail between the obligation on Tamra to attend to individual learners, the national curriculum goals, institutional goals, and enabling mathematical exploration. Tamra's choices ultimately mean that the obligation to mathematics suffers, yet they reflect a relevant practical rationality.

The teacher as a dilemma manager

Teachers make decisions, and decisions are informed by weighing up alternatives based on a range of factors. Previous research on teacher decision may be grouped, I suggest, into four non-disjunctive categories of factors informing the decisions. One group of studies focuses on teachers' beliefs and knowledge; for instance, Blackley et al. (2021) studied the relationship between experiences of affect, cognitive load, and awareness of available options. Another group of studies focuses on teacher noticing; for example, participating in a video club on noticing increased teachers' ability to use what they noticed about learner thinking to inform decisions (Wallin & Amador, 2019). A more recent group of studies concerns norms and patterns of participation established from previous experiences such as own schooling or teacher education; for example, the teacher education of two student teachers did not challenge their patterns of participation from earlier schooling (Ebbelind, 2020). Finally, a group of studies has engaged the obligations that exercise a pull on teachers, which may manifest as different and not always compatible goals for teachers: Thomas and Yoon (2014) describe how a teacher resolves conflicts between goals and the relation to the decisions the teacher makes. The teacher's ”conflict between [...] two goals was largely caused by the constraints of time, curriculum, and

Iben Maj Christiansen
Stockholm University and Mälardalen University

assessment, but also his desire to respect the cultural influences on his students' learning" (p. 232). Thus, the last group of studies hone more in on the practical rationality (see below) of teachers and how it plays out within the opportunities and constraints of their classrooms, regarding decisions to be shaped as much by institutional factors as by personal ones. Considering external factors is important for understanding why teachers make decisions that appear to go against their proclaimed views, and thereby the process of making changes to teaching practices. Managing external factors, the teacher is a "dilemma manager, a broker of contradictory interests" (Lampert, 1985), or in the terminology of practical rationality, a broker of professional obligations which may not pull in the same direction (Herbst & Chazan, 2011).

The experienced external demands and professional obligations of teachers affect the choices they make in their teaching. In a medium-scale study of university mathematics instructors, professional obligations helped explain misalignment between beliefs and practices (Shultz, 2022). Case studies of teachers from different contexts suggested that teachers may choose not to address a mathematically rich response from a learner out of consideration for the classroom community or may reduce learner agency to reach closure by the end of the lesson (Christiansen et al., 2023). And responses to a breaching experiment with 360 secondary mathematics teachers indicated that their willingness to break norms depended on both the norm and the professional obligations at stake (Erickson et al., 2021). When norms are broken, teachers can engage in negotiation of the didactical contracts that apply to the new situation, or they can default to a known instructional situation (Herbst & Chazan, 2012).

As the case discussed in this paper demonstrates, unpacking a negotiation of norms in light of the constraints on a teacher trying to change their practice, provides insights which ultimately may be brought to work in assisting or preparing teachers to handle these complex processes. The research question guiding the study is: *In which implicit negotiations of professional obligations does a teacher engage when she attempts to implement an exploration-requiring activity?*

The teacher in the study, Tamra, wants to change her teaching towards engaging learners in exploration activity, and the lesson analysed is her first attempt at doing so. However, as the lesson progresses, Tamra moves away from the exploration activity. To understand why, interviews with Tamra were analysed for indications of professional obligations and potential tensions between them.

Frameworks

Two theoretical perspectives have been utilised in this study: (i) for the analysis of Tamra's classroom teaching, I have drawn on notions within the theory of commognition; (ii) the reasons for Tamra's choices were investigated drawing on the notion of professional obligations.

In the commognitive theoretisation of Anna Sfard (e.g. 2008, 2016), one component of the mathematical discourse is routines. Explorations are discursive routines that lead to the production of substantiated narratives (Lavie et al., 2019). Rituals, on the other hand consist largely of ”imitating someone else’s former performance” (Nachlieli & Tabach, 2019, p. 255) ”for the sake of social rewards or in an attempt to avoid a punishment” (Lavie et al., 2019, p. 166).

What is a ritual to one learner, may be an exploration to another, though an observer is unable to tell the difference (Lavie et al., 2019). This makes the concepts inapplicable for analysing classroom activities, and in particular for analysing teaching. Instead, the *opportunities to learn* (OTL) – to engage in rituals or explorations made available through the teaching – can be considered, Nachlieli and Tabach (2019) suggest. Since learners may choose to engage in an exploration even if the relevant part of the lesson invites a ritual, Nachlieli and Tabach refer to such situations as *ritual-enabling OTLs*, defined as ”teachers’ actions that provide students with tasks that could be successfully performed by rigid application of a procedure that had been previously learned.” (Nachlieli & Tabach, 2019, p. 257). If, on the other hand, learners are not provided with a ritual to imitate, the lesson would require them to engage in an exploration, wherefore it is referred to as an *exploration-requiring OTL*, defined as ”teachers’ actions that provide students with tasks that could not be successfully solved by performing a ritual” (Nachlieli & Tabach, 2019, p. 257).

Learners often have to learn a routine as a ritual and gradually move from imitating the routine to focusing on how to reach the desired outcome, that is, treating the routine as a narrative; this process is referred to as deritualisation (Lavie et al., 2019, in passim). This shift in discourse happens when learners take part in the discourse while exerting effort to shift attention to the outcomes of the routine (Sfard, 2020). Deritualisation is characterised by increased flexibility, bondedness, applicability, performer agentivity, objectification, and substantiability (Lavie et al., 2018).

In the analysis of Tamra’s lesson, I focused on the three latter of the characteristics of deritualisation, from a teaching perspective. This means that instead of focusing on the characteristics of deritualisation evident in the discourse of the learners, I analysed Tamra’s interactions with the learners for invitations for learner agentivity and substantiability, as well as the extent to which her discourse utilised mathematical objects. This allowed me to determine the extent to which Tamra implemented an exploration-requiring OTL (see Christiansen et al., 2023 for details on this approach).

The analysis of Tamra’s lesson served as a backdrop for identifying the choices she made as the lesson progressed and what informed them. For this purpose, interviews with Tamra were analysed for references to professional obligations. Professional obligations are generated by the ”hold of the environment on the position of the teacher” and Herbst and Chazan (2011; 2012) categorise these according to their source: the individual learner, the socio-cultural world, the institution, and the discipline. The obligations shape the teacher–learner relationship and operate on the norms of the teaching.

The teacher-learner relationship exists because the learners must acquire knowledge which the teacher has (or, within a participatory view on learning, because the learners are expected to increase their participation in the discourse of which the teacher is a more adept participant). But since this learning may not happen, the knowledge is ”at stake” (Herbst & Chazan, 2011; 2012). The teacher must enable learners' mathematical work that can be ”exchanged” for the knowledge at stake. Often, a particular norm has been established for the mathematics teaching in each classroom, and changing this norm requires renewed negotiation of practices and relationships. However, it is also subject to the constraints of the institution, the curriculum, and mathematics itself. Therefore, negotiating the norms of the classroom takes place not just between teacher and learners but in relation to the professional obligations.

Obligations linked to *individual* learners can, for instance, be to consider the unique strengths and needs of each learner, even as these are represented by their legal guardians (Herbst & Chazan, 2012; Schultz, 2022). The teacher is also expected to represent the *discipline* as it is understood in the institutions of society. Teaching is always taking place within a broader socio-cultural context with values, norms, and practices, and this constitutes *interpersonal* obligations to create a socio-culturally appropriate classroom environment (Herbst & Chazan, 2012; Schultz, 2022). Finally, there are *institutional* obligations on teachers from curricula, policies, school rules, etc., which include fixed lesson times, rules about use of space, and so forth.

Methods

Tamra is a participant in the Swedish TRACE project which follows novice mathematics teachers in the first years of their profession. In the project, eleven of her lessons were video-recorded, starting from two lessons during her teaching practice in 2018, and biannually with a one-year interruption due to Covid-19 restrictions, until May of 2022. Five interviews with Tamra were conducted: one about her impressions of her teacher education, three about a sequence of lessons observed, and one looking back over the period and her development. Lessons were video-recorded with a focus on Tamra's, not learners' activity, and interviews were audio-recorded. This study discusses Tamra's first experiment with an exploration-enabling OTL, a lesson on prime numbers.

For this study, I conducted two analyses. The first analysis was of the lesson, and it allowed me to identify that the lesson moved from offering exploration-requiring OTLs, in line with Tamra's intentions, to offering ritual-enabling OTLs. The lesson was first divided into units of analysis, where each unit started with the initiation of a task by the teacher and ending with the closing of the task (following Nachlieli & Tabach, 2019). Within these units, Tamra's utterances were coded for three indicators of exploration-enabling versus ritual-enabling OTLs (following Christiansen et al., 2023). Because she generally referred to mathematical objects in her discourse, the analysis here focuses on the extent to which Tamra invited learner agentivity and learner substantiations.

I coded for overtures for learner agency by the openness of the teacher's questions and invites for participation to learners (see also Sfard, 2016). Invitations for substantiations were coded by whether Tamra encouraged learners to reproduce substantiations or follow steps of a procedure, alternatively encouraged them to generate their own substantiations. Frequent overtures for learner agentivity combined with encouragement of learners' own substantiations were seen as indicative of exploration-requiring OTLs, while the opposite codes were seen as indicative of ritual-enabling OTLs (Christiansen et al., 2023 discuss hybrids between these extremes).

The second analysis was a content analysis of the interviews with Tamra, identifying references to professional obligations. Since obligations are characterised as a "hold of the environment on the position of the teacher", I used content analysis to identify instances where Tamra referred to factors restricting her teaching in any way. Next, I coded these obligations with reference to (a) individual learners or their parents, (b) the class environment, (c) mathematics, (d) institutional factors such as the curriculum, lesson time, or school rules, and (e) other. I also noted when Tamra referred to norms or beliefs she had to confront or change in the process of adjusting her teaching.

In this paper, I focus on the obligations Tamra experienced, but I relate these to how the lesson progressed. Therefore, the next section presents the lesson and my analysis thereof, as a backdrop to the results section.

Tamra's lesson on prime numbers

When asked what she wanted to achieve with the lesson, Tamra said that she wanted learners to identify prime numbers (a new mathematical object used in an exploration), factorise (a new routine), use divisibility rules (utilise previous routines in explorations), cement concepts of even and odd, and use the even/odd distinction in checking for divisibility (utilise previous routines).

The lesson had three distinct parts, as well as an almost eleven-minute period spent on organising groups, about which Tamra later expressed frustration. The first part of the lesson was organised as whole-class teaching, with Tamra stating the definition of prime numbers after which learners identified small prime numbers by inspection. The second part of the lesson was also whole-class teaching, where the teacher guided learners through factorisations of 24. The last 19 minutes of the lesson constituted the announced exploration, and my analysis confirmed that it started out as an exploration-requiring OTL. "Can you

Table 1. *Tamra's lesson on prime numbers and coding of overtures for learner agency (open/closed) and invited substantiations (learner generated/given)*

Time	Focus	Overtures for learner agency	Invited substantiations
02:02–09:41	Whole-class interaction, with learners being tasked with identifying whether 2, 3, 4, 5, … 10 are prime or composite. At the end of this part, Tamra refers to previously learned rules for divisibility.	Open and closed questions interspersed.	Learners are asked to explain their answers.
09:42–14:44	Learners are asked to identify whether 24 is prime or composite, but because learners suggest different initial factorisations (4x6, 2x12, 3x8), this evolves into three subroutines of factorising and a question about similarities between factorisations.	Open and closed questions interspersed.	Learners are asked to explain their answers.
14:45–25:29	Dividing learners into groups and distributing paper and coloured pens.	None.	None.
25:30–44:30	Learners work in groups deciding whether the natural numbers in different intervals (e.g. 200–250) are odd, even, prime, and/or composite. The class gets rowdy and has to be calmed before the nine-minute report back at the end of the lesson.	First few questions are open, the remainder closed.	First invitations are for learners' substantiations, the remainer requests to follow Tamra's procedure.

find factors?", Tamra asked repeatedly [1]. There is no indication that she wanted learners to follow a particular procedure.

Shortly after the groups started to work, while Tamra circulated, the lesson took a turn. In one group, Tamra engaged the learners as she had previously in the lesson: "Are there two numbers we can multiply to get thirteen?" However, the next group she approached was exploring larger numbers, and Tamra instead asked whether the number in question at the time could be divided by 2, 3, or 5. She also used this to identify an even number which the learners had categorised as prime. She used the same strategy with the next group she visited, where learners were struggling to categorise 151. She did not go back to asking the more open question – whether learners could find factors.

The routine of checking for divisibility by 2, 3, and 5 replaced the previous approach of inspection by recognition (which failed the learners working with numbers larger than 100 and learners who did not recognise products from the multiplication tables below 100). In the report-back session from the groups, Tamra verified or rejected the prime numbers suggested by learners only by inspecting for divisibility by 2, 3, and 5.

Insisting on a particular procedure, taking control over when the task is complete, and using closed questions, changed the situation from an exploration-requiring to a ritual-enabling OTL. It is easy to find fault with Tamra's actions, her decision to take control over the choice of procedure and the closing conditions of the task. However,

> we could think of ”error” in instruction – really teaching that deviates from what might be deemed desirable – not as an indication of misfit, ill will, or lack of knowledge, on the part of the practitioner. Rather, we should think of this ”error” as an indication of the possible presence of some knowledge, knowledge of what to do, which is subject to a practical rationality that justifies it. (Herbst & Chazan, 2011, pp. 428–9)

What made Tamra change the exploration in which she had so wanted the learners to engage? To answer this, I turn to the results of my analysis of the interviews with Tamra.

Results – obligations experienced by Tamra

The view on learners and mathematics teaching which Tamra said dominates at her school is to focus on the ”weaker” learners, reduce the number of at-risk learners, and give everyone the opportunity to pass. This is in line with the expectations on teachers in Sweden generally, and so was coded as an *interpersonal* obligation. It is not unrelated to the socio-economic context of the school – Tamra deals with learners who have given up on school, learners with challenging home situations, and even drug abuse.

Tamra also talks about the pressure she feels from some parents to pay special attention to their child and help the child achieve well. This obligation towards *individual* learners, and in particular those who are identified as at-risk, creates tensions for Tamra when it meets the *institutional* obligation to teach to and ”cover” the curriculum. When asked about the challenges she experiences, she starts listing these:

Tamra: To adapt the teaching so that it suits and reaches everyone. With which pacing must I teach that fits with the content and plan of the curriculum? The greatest challenges so far have been how I can engage learners with drug problems and for whom school has no importance; I still work on that to help the learners.

Tamra wants the learners to engage in an exploration leading to separating positive integers below 250 into prime and composite numbers (an obligation to *mathematics*). At the same time, she is under obligation, from the curriculum (*institutional*), to ensure that each learner can utilise the divisibility rules for 2, 3, 4, 5, 6, 8, 9, and 10 which they have learned in a previous lesson, and she expects the learners to bring that into play in the exploration. However, this was not made clear to the learners; there was no negotiation of task in this – for Tamra and her learners – novel organisation of the classroom activity. At the same time, the disallowance of calculators may facilitate the relevance of using divisibility rules, but it makes it harder for learners to quickly test a hypothesis about a number – such as whether 151 is divisible by seven or not.

Retrospectively, Tamra acknowledges that the many learning objectives for the lesson were counterproductive, particularly for a class where many learners dislike or have given up on mathematics (obligations to the *individual*).

Tamra: The thought was that maybe they can cement their knowledge about even and odd numbers and that [the exploration] could help them apply divisibility rules. But it was too much for such a group, which contains many [learners] who do not have strong foundations in their mathematics and have short concentration spans.

Tamra had 45 minutes for the lesson. She needed to first introduce the definition of prime and composite numbers and ensure that learners understood. In this process, she got into different factorisations of 24, and an extended period of time was spent organising the group work. The group work progressed very slowly, and because Tamra wanted a report-back/sharing phase at the end, only 10 minutes were left for the exploration which she had declared to be the main focus. The time frame is part of the *institutional* obligation to which Tamra must adhere. Yet she addressed different ways of factorising because she wanted to pick up on learners' questions.

Tamra: I did it just because [different factorisations] occurred to the learners, as I have an ability to make spontaneous changes during the time of the lesson based on what learners ask, and sometimes their questions awaken ideas which lead to a better result.

This is one way in which Tamra handles the *interpersonal* obligation towards ensuring fruitful conversations with and between learners; but in this case, it is at tension with the institutional obligation to complete the lesson within 45 minutes, because tomorrow's lesson must be about something else – as Tamra laments after the lesson. In a way, Tamra's norm of adjusting the lesson when constructive opportunities arise for engaging a mathematical narrative beyond the lesson objectives (here the fundamental theorem of arithmetic) clashes with her intention of giving learners time for exploration. It is hence no wonder she chooses to "*default to an instructional situation*, namely by framing the exchange according to norms that have framed other exchanges" (Herbst & Chazan, 2012, p. 606) – in this case by introducing a procedure.

Discussion

Tamra had two main objectives for the lesson – to engage learners in an exploration of the mathematical objects of prime and composite numbers, and that they should do so using previously introduced routines for checking divisibility. It appears that mathematics is sacrificed to carry the lesson through to completion, in two ways. First, the exploration-requiring OTL is transformed into a ritual-enabling OTL, reflecting a particular way for learners to engage mathematically. Second, the procedure of checking for divisibility of only 2, 3, and 5 has the potential to generate incorrect categorisation of composite numbers as prime (though it does not happen in the lesson).

In the paper, I tried to unpack Tamra's reasons for her instructional choices in the lesson using the lens of professional rationality. The two objectives are not in themselves in conflict. However, they become so – as in the study of Thomas and Yoon (2014) –because of the institutional constraints of time, together with the obligation to individual learners and the interpersonal obligations of not letting learners fail. Through identifying the obligations that reflect ”the hold of the environment” on Tamra as a teacher, I have aimed to map the borders of the space in which she can exercise her rationality. My analysis shows how Tamra's own views and efficacy, the long-term desired outcomes of teaching vis-à-vis the curriculum, the learners, and the institutional routines create tensions which inform the choices Tamra could and did make. This is in line with previous research. What is unique about Tamra's case, however, is that she is an early career teacher deliberately trying to change her teaching.

The lesson can be seen not just as a didactical situation between Tamra and her learners in a particular instance of space and time, but as part of Tamra's journey to develop her teaching. While Tamra's norms for teaching are still being established, the lesson constitutes a personal breaching experiment (the idea of which is often ascribed to Garfinkel, e.g. 1964). By immersing herself in an instance of a practice that alters her usual teaching, she provides herself – and the research team – with an opportunity to see how she repairs the breaches with the norms she deliberately makes.

As shown above, when confronted with tensions, Tamra chooses to default to a more ritual-enabling OTL. She started by explaining to learners that this was going to be a different lesson, but as this description does not clarify the different activity expected from learners, it does not constitute a negotiation of the didactical contract, as proposed by Herbst and Chazan (2012). In the interview about the lesson, Tamra feels that her instructions to learners should have been clearer, but this does not capture the difference between explorative and ritual routines.

Tamra: I think that we teachers choose and through the choices we determine what the classroom environment becomes. [...]So the instruction was lacking a bit, my instructions to them, besides that it was, no, more was needed, it should have been well planned, it wasn't well planned. It was planned but not really thought-through planned.

Still, Tamra does not just see this as her shortcoming; she is aware of the expectations, the hold of the environment, on her. In her own words, she must be a mathematics teacher, a special needs teacher, a psychologist, and a researcher/ practice developer.

What can the case offer mathematics education research, teachers, and teacher educators? It reminds researchers to keep in mind the constraints under which teachers operate and illustrates the relevance of the notion of practical rationality in doing so. Did she indeed sacrifice the mathematics? Is that the best way to describe her choice? Or did she make a choice which enabled her to ”repair” the deviation both from her normal practice and from her intended lesson, so that she could draw the lesson to a close, sufficiently cover a part of the curriculum, and leave the majority of the learners with some – even if incomplete – idea of what prime numbers are and how to identify them? After all, there is a day after this one, and a year after this one, in which it is up to Tamra – and her learners – to correct, adjust, and expand concepts; and in which Tamra can work to correct, adjust, and expand her teaching.

Teachers and teacher educators can utilise the discursive concepts of commognition to more precisely describe different opportunities to learn and what this means to the activity expected from learners. Together with the notion of the didactical contract and negotiation of the same, this becomes more than an analytical tool, it becomes a practical tool to utilise when wanting to implement exploration-enabling OTLs. Instead of despairing at the discrepancy between visions of good teaching and practice, defaulting to deliberately chosen aspects of instructional situations and negotiation of the didactical contract can move practice in desired directions. An awareness of the professional obligations which positions the teacher as a broker may assist in conscious decision-making. And knowledge of the dimensions of deritualisation and related teaching move can guide the negotiations of the didactical contract as well as inform instructional choices. There is much for teacher education to work with here.

Acknowledgments

Thank you to Tamra, colleagues, and reviewers. This study is part of the TRACE project supported by the Swedish National Research Foundation, grant number 017–03614.

References

Blackley, C., Redmond, P. & Peel, K. (2021). Teacher decision-making in the classroom: the influence of cognitive load and teacher affect. J*ournal of Education for Teaching*, 47 (4), 548–561. doi: 10.1080/02607476.2021.1902748

Christiansen, I. M., Corriveau, C. & Pettersson, K. (2023). Hybrids between rituals and explorative routines: opportunities to learn through guided and recreated exploration. *Educational Studies in Mathematics*, 112 (1), 49–72. doi: 10.1007/s10649-022-10167-z

Ebbelind, A. (2020). *Becoming recognised as mathematically proficient: the role of a primary school teacher education programme* [Ph D thesis]. Linnaeus University.

Erickson, A., Herbst, P., Ko, I. & Dimmel, J. (2021). When what routinely happens conflicts with what ought to be done: a scenario-based assessment of secondary mathematics teachers' decisions. *Research in Mathematics Education*, 23 (2), 188–207. doi: 10.1080/14794802.2020.1855600

Garfinkel, H. (1964). Studies of the routine grounds of everyday activities. *Social problems*, 11 (3), 225–250.

Herbst, P. & Chazan, D. (2011). Research on practical rationality: studying the justification of actions in mathematics teaching. *The Mathematics Enthusiast*, 8 (3), 405–462. doi: 10.54870/1551-3440.1225

Herbst, P. & Chazan, D. (2012). On the instructional triangle and sources of justification for actions in mathematics teaching. *ZDM*, 44 (5), 601–612. doi: 10.1007/s11858-012-0438-6

Lampert, M. (1985). How do teachers manage to teach? *Harvard Educational Review*, 55 (2), 178–194.

Lavie, I., Steiner, A. & Sfard, A. (2018). Routines we live by: from ritual to exploration. *Educational Studies in Mathematics*, 101. doi: 10.1007/s10649-018-9817-4.

Nachlieli, T. & Tabach, M. (2019). Ritual-enabling opportunities-to-learn in mathematics classrooms. *Educational Studies in Mathematics*, 101 (2), 253–271. doi: 10.1007/s10649-018-9848-x

Shultz, M. C. (2022). The rationality of undergraduate mathematics instructors: the choice to use inquiry-oriented instructional practices. *Journal for Research in Mathematics Education*, 53 (3), 227–246. doi: 10.5951/jresematheduc-2020-0306

Sfard, A. (2008). *Thinking as communicating: human development, the growth of discourses, and mathematizing*. Cambridge University Press.

Sfard, A. (2016). Ritual for ritual, exploration for exploration. In J. Adler & A. Sfard (Eds.). *Research for educational change: transforming researchers' insights into improvement in mathematics teaching and learning* (pp. 41–63), Routledge.

Thomas, M. & Yoon, C. (2014). The impact of conflicting goals on mathematical teaching decisions. *Journal of Mathematics Teacher Education*, 17, 227-243. doi: 10.1007/s10857-013-9241-8

Wallin, A. J. & Amador, J. M. (2019). Supporting secondary rural teachers' development of noticing and pedagogical design capacity through video clubs. *Journal of Mathematics Teacher Education*, 22 (5), 515–540. doi: 10.1007/s10857-018-9397-3

Note

1 All quotes from the lesson and the interviews were transcribed in Swedish and translated by me.

Designing teaching for creative mathematical reasoning combined with retrieval practice

Jan Olsson, Maria Larsson, Carina Granberg, Ann-Sofie Persberg, Lena Salomonsson and Denice D’Arcy

The aim of the study was to explore teacher support to students’ mathematical reasoning when they retrieve and use prior knowledge to learn new mathematical content. Two lessons were designed, one with the purpose to engage students in retrieving prior knowledge of angles and shares, and the following to learn how to construct pie-charts. The teacher’s interactions with students were recorded and the result showed that the teacher could support students’ reasoning by showing interest in their thinking and asking challenging questions regarding clarification and justification of their mathematical reasoning.

During the last decades, research in mathematics education have approached an agreement of that students, in substantial parts of teaching, should be active in solving and justifying solutions of mathematical problems (see e.g. Schoenfeld & Sloane, 2016; Lester & Cai, 2016). However, there are several issues to consider if such a teaching approach would be successful in terms of sustainable learning. Criticism directed to approaches of learning through problem solving often concerns students’ ability to handle a large amount of information in working memory (see e.g. Kirschner et al., 2006; Sweller, 2020). Solving mathematical problems is complex and demands the solver to use prior knowledge to construct and assess the plausibility of the solution. If working memory is overloaded no learning will take place (Sweller, 2020). On the other hand, there is research showing that students who practice tasks that require construction of the solution learn better compared to those who solve the same tasks, but are provided with methods (see e.g. Jonsson et al., 2014). However, students who are not successful in constructing useful methods during practice do not learn what was intended (Olsson & Granberg, 2019). Hence, there is an opportunity to increase learning outcomes by appropriate teacher assistance to students’ problem solving, provided that the teacher does not reveal a solution method.

Jan Olsson, Umeå University, Mälardalen University and Dalarna University
Maria Larsson, Mälardalen University
Carina Granberg, Umeå University
Ann-Sofie Persberg, Lerbergsskolan
Lena Salomonsson, Lerbergsskolan
Denice D’Arcy, Lerbergsskolan

Creative mathematical reasoning and retrieval practice

A way to understand what support students need is to pay attention to their reasoning. Lithner (2008) has defined reasoning in which students construct their own solution method and formulate arguments for the solutions as *Creative mathematical reasoning* (CMR). We have for several years investigated how a teacher can support students' CMR (see e.g. Olsson & D'Arcy, 2022; Olsson & Granberg, 2022) and recurrently noticed that students often need assistance to retrieve the prior knowledge required to solve the current task. A concept of retrieving knowledge from long-term memory is *Retrieval practice* (RP) which in studies has shown to strengthen learning and memory (Dunlosky et al., 2013). Theoretically, the retrieval activity creates stronger associations among the individual's prior knowledge (ibid.). Hence, since learning mathematics by CMR has shown creating strong memory, CMR may enhance RP. On the other hand, RP may strengthen the ability to retrieve necessary knowledge to engage in CMR. Therefore, a combination of CMR and RP is of interest when designing mathematical teaching.

In this study, problems are defined as tasks for which students do not have a pre-existing method for solving the problem, i.e., they need to construct (parts of) the method themselves. Lithner (2008) describes the cognitive process of constructing methods and providing mathematical founded arguments supporting that method as CMR. A series of studies have shown that students who during interventions engaged in CMR to solve problems outperform, on post-tests, students who were using provided methods. (e.g. Jonsson et al., 2014; Norqvist et al., 2019). Recent studies have looked further into learning outcome depending on practicing on CMR-tasks or on tasks that provide a solution method. Jonsson et al. (2020) conducted a series of experiments, which showed that students who solved CMR-tasks (constructing methods) gained a better mathematical understanding, which they could use during post-tests to solve transfer tasks. Moreover, Norqvist et al. (2019) found that students who engaged in CMR tasks were more attentive to the mathematics embedded in the problem compared to those who were provided with a method. That is, constructing a method is likely to promote a deeper mathematical understanding compared to merely implementing given methods. Furthermore, Olsson and Granberg (2019) found in their study that knowledge gained through CMR resulted in better memory consolidation and a more comprehensive mathematical understanding. During a post-test, these students could apply what they practiced to solve problems more effectively compared to them who practiced using provided methods. The aim of this study is to explore teacher support to combine CMR and RP. The context is two planned lessons, one where students retrieve usable prior knowledge and the following where students with aid of that prior knowledge learn new mathematical content for them. The research questions are: (1) *How can a teacher assist students to retrieve prior knowledge?* and

(2) *In what ways can a teacher support students to utilize the prior knowledge they retrieved in a preceding lesson when learning new mathematical content?*

Framework

Theoretically, the teaching design is based on the definitions of CMR and RP. These concepts are operationalized by principles of teaching for CMR developed in our earlier studies (Olsson & D'Arcy, 2022; Olsson & Granberg, 2022).

The CMR framework emphasizes the active process of creating and constructing knowledge, particularly students' ability to create their own solutions of tasks and formulate mathematical arguments for the solution. A CMR-task does not include a method to employ. The RP framework, on the other hand, emphasizes memory consolidation through the act of actively recalling information from memory. Theoretically, through the active process of retrieving knowledge from long-term memory, RP starts a process of re-consolidation which creates stronger connections between memorized items.

Our earlier studies focus on developing principles for teaching where students learn through CMR (Olsson & D'Arcy, 2022; Olsson & Granberg, 2022). The principle for CMR-teaching is: *If the goal is that students will learn mathematics by CMR, tasks and teacher-student interactions should guide students into initiating, developing, justifying, and verifying their reasoning when constructing and formulating arguments for the solution*. This study aims at exploring possibilities for combining CMR and RP and the proposition (initial principle) for RP is: *If students are supposed to use prior knowledge when learning new mathematical content, students should, in a separate lesson before learning new mathematics, have opportunities to activate adequate prior knowledge*. These principles will guide the intervention's design and constitute a base for outcome conjectures, which in turn will guide the analysis.

Method

The method is based on the design of an intervention guided by the earlier presented principles for CMR and RP. Tasks and teacher support were prepared with the intention to engage students in CMR, and the students' prior knowledge was considered. Conjectures were made regarding how the tasks would engage students in constructing the solutions and how the teacher could support initiating, developing, justifying, and verifying reasoning. These conjectures were foci for the analysis.

Methodology

The current paper reports from an intervention within a larger project, in which a teaching design is developed and investigated. The innovative character of the study means it is not likely that regular classrooms will meet the conditions the investigations require. Therefore, teaching aiming at promoting CMR must be designed (see Cobb et al., 2016). Design-based research methodology combines the construction of scientific knowledge and its applicability in educational contexts (Cobb et al., 2016; Schoenfeld, 2007). A fundamental issue of design research is that principles guide rather than determine a design and therefore require creative input, and development through feedback on trials (Swan, 2008). Typically, a design-based project is carried out in iterative interventions, each intervention starting in assumptions of the outcome. The design of a specific intervention (as in this paper) is based on principles developed in previous interventions within a project. Analysis of outcomes of the specific intervention guides developments and revisions of the principles (Cobb et al., 2016).

Participants

22 students aged 13–14 years agreed to participate in the study. Their parents were informed and signed agreements for the students to participate. The teacher considers the class to be on an average level in terms of mathematical ability and there are both high and low achievers. The teacher is experienced and participates in a long-term project this study is a part of. This is her first attempt to teach according to the design.

Intervention design

The intervention consisted of two lessons. The overall learning goal was to construct pie-charts and it was considered the students could use prior knowledge of angles, fractions, decimal numbers, and percent. The first lesson aimed at activating that prior knowledge. The goal was that all students would retrieve definitions of angles, how to measure angles, the sum of angles in rectangles and triangles, that a full circle is 360° and how to calculate percent and shares. The tasks both had the character of retrieving factual knowledge (e.g. how many degrees is a full circle?) and explaining concepts (e.g. explain why 1/4 is 25 %). The teacher prepared in collaboration with the researchers how to interact with the students. Conjectures were made that solving the tasks in pairs would engage students in remembering at least parts of properties such as acute angle, right angle, obtuse angle, and abilities to measure and construct angles by ruler and protractor. Furthermore, if parts were not remembered the students were anticipated to, in collaboration with peers and the teacher, retrieve those parts by CMR.

The researchers created, together with the teacher, conjectures on how teacher-student interactions that do not involve facts and procedures for how to solve the task can promote students to retrieve relevant knowledge via RP (knowledge that the student needed to engage in CMR). Once the students had retrieved prior knowledge, they would work together to formulate arguments for how the retrieved knowledge is correct. In cases where the students do not try to retrieve prior knowledge, the teacher needs to challenge them to do so. Furthermore, the teacher should not verify students' conclusions as correct or not, but instead ask questions like *how can you be sure you are right?* and prompts like *discuss with your peers*. Considering that it was the teacher's first attempt; the conjectures were deliberately few and assessed as reasonable to implement.

The goal of the second lesson was for the students to learn how to construct a pie chart and to do that they needed to find methods to translate percentages into degrees. The students were given three sub tasks. In the first, students could draw simple pie charts without translating percentages to degrees (e.g. 50 % and 25 %). The second task was slightly more difficult, but students were able to draw pie charts by making estimates (e.g. 20 %). The third task was designed so that students would see the need to translate % into degrees (e.g. 17 %). Conjectures for the second lesson were that all students would promptly construct the charts of 50 % and 25 %. Thereafter, they would estimate the chart of 20 %. When finding the chart of 17 % difficult, they would engage in reasoning how to, based on the knowledge that a full pie-chart is both 100 % and 360°, calculate how many degrees are 17 %. Conjectures on how teacher-student interactions would support students were to not verify correct or incorrect answers but rather ask the students how they could be sure whether their answers were correct or not. Difficulties anticipated were to translate percent into degrees. Here the teacher could encourage them to use the knowledge that a full circle is 360°, and if necessary, ask what percentage a full circle represents.

The structure of the lessons and data collection

The teacher had a brief introduction informing the students they would work on tasks in pairs. During the first lesson, the students were told they had met the mathematical content before and for the second that they would have a chance to learn how to construct pie charts. The students in both lessons were informed that they could ask the teacher for help whenever they needed, and they presented the solutions on separate papers. The teacher and a researcher carried a voice-recording device each. The teacher had the routine to in every interaction mention the students' names in purpose to facilitate pairing of recordings from both lessons. The researcher followed the teacher and after a teacher-student interaction the researcher stayed a while to catch reasoning between students after interacting with the teacher. Thereafter the researcher caught up with the teacher. The students were told that they could not expect any help from the researcher. However, the researcher could ask questions aiming at clarifying parts in students' reasoning. Hence, data consisted of recordings of teacher-student interactions, students reasoning after the interactions, and students' written solutions.

Analysis method

The voice-recordings were transcribed verbatim. Teacher-student interactions were then identified where the students either recalled prior knowledge, related this to new knowledge or gave explanations thereof. Then, a first step was to characterize the interactions based on the contributions from students and the teacher, respectively. Patterns in the interactions were searched for and categories were formed. Thereafter, based on conjectures of how the design would work, it was described how the teacher aided students to retrieve prior knowledge (lesson 1) and supported students to utilize prior knowledge when learning how to construct pie charts (lesson 2).

Results

During lesson 1, the retrieval lesson, 23 teacher-student interactions were recorded. In five of them the students had solved a task (or several tasks) correctly and explained their solutions for the teacher. In the other 18, the students needed teacher support. Initially, the students rather easy retrieved factual knowledge such as labels of different angles and procedures like measuring acute angles. Some arguments, typically on request from the teacher, were expressed. Most students succeeded in measuring acute angles while some discussions between students and the teacher could be observed when measuring obtuse angles. However, difficulties appeared when measuring angles larger than 180°.

Student C: This one is not 50°. This is not the outside. Is it this one you must measure?

Teacher: You do not know which one to measure. How did you do in the earlier ones? Why don't you know now?

Student D: This one is in the wrong direction. It is on the outside. Outside the [small] angle.

Teacher: Okay, how do you do then?

Student D: Well, the small angle (students work in silence).

Teacher: Ah, you are measuring the small angle. Why do you subtract from 360°?

Student D: Because that is the whole [circle].

The teacher, typically, did not verify whether the students were correct or not. The questions support the students to use knowledge on how to measure angles smaller than 180° in reasoning how to with the aid of a 180°-protractor calculate an angle larger than 180°. What was not obvious but reasonable is that the students realized that if they know the smaller angle, they could use the knowledge that a whole circle is 360° to calculate the larger angle. This is an example where the teacher could support students to retrieve knowledge both as fact and through reasoning without providing facts or procedures.

The second lesson on learning how to construct pie charts included 18 teacher-student interactions. Five of them could be categorized as the teacher supported students to retrieve content from the retrieval lesson and to use that knowledge to solve the current tasks. In five interactions the students managed to retrieve knowledge from the retrieval lesson but needed teacher support to solve the current tasks. In eight interactions the students had successfully solved tasks and explained them for the teacher. Initially the lesson followed the anticipated path, the students promptly constructed pie charts with a 50 % pie and with a 25 % pie. Some differences in the approach of the 20 % and 80 % chart were observed. Most students recognized that 20 % is 1/5 and estimated how to create a corresponding pie in the chart. Some students successfully realized that 1/5 of 360° is 72° and in the third task, constructing a chart with a 17 % pie, they replicated the approach. The following is an example where the teacher supported students to retrieve and use knowledge from the previous lesson. Student O and F had without problems constructed pies of 50 % and 25 %. Approaching the task to construct a pie of 20 %, they initially retrieved knowledge that 20 % is 1/5, which they used to estimate a pie. However, when the teacher asked them how they could be sure that the pie was 20 %, they could not answer. The teacher asked them if they could measure how many degrees the pie constitutes.

Student O: I don't know how to do that.

Teacher: Do you remember what we did yesterday? You constructed and measured angles.

Student F: [student measures the angle] 45°

Teacher: Okay, is that correct? 45° is 20 %? Can you check that?

Students: [silence]

Teacher: Is that correct, that 45° is 1/5? What is 1/5?

Student O: 1/5 of 360.

Teacher: Try that [leaves the students]

This is an example of an interaction when the teacher supported the students to retrieve content from the earlier lesson. Later the teacher returns to students O and F.

Teacher: Okay, tell me what you have done.

Student F: This is 17 % and I need to know what that is in degrees. So, I divided 360 by 100 to know how many degrees is 1 %. Then I multiplied 3.6 with 17 and that is 61,2 which is approximately 61. Then I constructed an angle of 61°.

In this sequence, the students have solved the task, reasonably aided by the support they earlier had from the teacher. Quite a few students did not interact with the teacher until they had solved all the tasks. The following is an example where the students had solved the task before interacting with the teacher, who typically asked them to explain what they had done and how they were thinking.

Student A: We divided 360 by 5.

Teacher: Why did you divide it by 5?

Student A: Because 20 % equals 1/5.

Teacher: In the other task [17 %] it seems like you had another strategy?

Student A: Yes, we divided by 100 and then multiplied with 17.

Teacher: And?

Student A: Then we had 61,2 which is approximately 61, and then we constructed the angle.

These students use content from the retrieval lesson with confidence. The teacher does not need to support the students' learning, but her questions allow the students to verify that they have the knowledge according to the lesson's learning goal. However, there were examples when students apparently had used content from the retrieval lesson but still had problems with the task asking them to construct a chart with a 17 % pie. The following is an example where the teacher supported students to solve the task with aid of the knowledge from the earlier lesson that the students have retrieved independently. Student H and I had successfully calculated the degrees for 20 % and constructed the angle. Now they reason that 17 % must be slightly smaller.

Teacher: Okay, but how can you do to get the exact size?

Student I: I think that the whole circle is 100 %. We can divide 360 by 100.

Student H: And then we can multiply with 17.

Teacher: So, you divide 360 by 100. What did you get?

Student I: 1 %, 3.6°.

Teacher: And then?

Student I: 3.6 times 17.

From this point the students promptly constructed a 61° angle in the chart. This is an example in which the students obviously used content from the retrieval lesson and needed minor support from the teacher to translate 17 % into degrees. Notably, the teacher support does not include direct instructions on how to calculate the angle, but a question that helped students to focus on the problem; How to transfer 17 % to a fraction. In summary, the teacher's strategy to start any interaction with asking the students to explain their thinking helped her to adjust further support. Students who had solved the tasks independently could be challenged to explain and justify their solutions. Students who in their reasoning revealed that they had retrieved necessary prior knowledge could be guided to focus on the challenging parts of new content, and students who did not use content from the retrieval lesson could be notified of that.

Discussion

This study asks how a teacher can support students to retrieve prior knowledge and how a teacher can support students to use the retrieved knowledge when learning new mathematical content. The result indicates that quite a few students need to reason in retrieving mathematical content not taught recently. In the first lesson most students in this sample were able to do that. A reason can be the teacher's approach to ask students to explain their thinking and give them time to do so. Furthermore, splitting the activities of activating prior knowledge and learning new content may help students to concentrate on retrieving earlier learned mathematics rather than try to understand new content and retrieving necessary knowledge simultaneously. From the teacher's perspective, it may also be easier to support students' knowledge retrieval in a separate lesson than when students need support to use prior knowledge when learning new mathematics. In our earlier studies a recurrent issue has been to support students' retrieval of prior knowledge when solving CMR-tasks aiming at learning new content (Olsson & Granberg, 2022; Olsson & D'Arcy, 2022).

In the second lesson, some students did not spontaneously use content from the retrieval lesson. However, after the teacher encouraged them to retrieve and use such content, there were examples when students promptly solved all tasks including the most challenging one with a 17 % pie. It is reasonable that even though the students needed support for the retrieval, knowledge from the first lesson could be retrieved more easily compared to if there had not been such a lesson. For the teacher, the knowledge that the students retrieved appropriate prior knowledge in the previous lesson may help formulating supporting questions. Both the first and the second lesson were designed based on the principle for CMR, which means the teacher encourages the students to initiate, develop, justify, and verify their reasoning. Experiments building on similar theories have been found increasing the ability to transfer knowledge into new tasks (Jonsson et al., 2020, Olsson & Granberg, 2019). In summary, in this intervention, the design seems to facilitate some of the known difficulties associated with learning new mathematical content through CMR. Anyhow, further research is needed, particularly the significance of the retrieval activities must be investigated.

The principle for CMR-teaching has not primarily been tested in this study. It has guided the design of the intervention and much speaks for that the conditions allowed students to initiate, develop, justify, and verify reasoning, which is a result replicated from other studies (e.g. Olsson & Granberg, 2022). However, the proposition for RP which guides the design of separating activation of prior knowledge from using prior knowledge when learning new content, may be evaluated based on the results. All students to some extent needed to reason when retrieving prior knowledge. Objections to teaching where students learn from problem solving often state that there is too much load on working memory both to retrieve prior knowledge and implement a solution strategy (Kirschner et al., 2006; Sweller, 2020). We consider that allowing students to activate prior knowledge in a separate lesson close to learning new content will facilitate the retrieval of necessary prior knowledge when solving problems aiming at learning new content. The results showed quite a few students who fluently or with minor support from the teacher retrieved knowledge from the first lesson when constructing pie charts. However, there were some students who in the second lesson needed extended teacher support to retrieve necessary prior knowledge. These students may benefit from an additional separate retrieval activity. Studies of RP have shown that repeated retrieval activities will further strengthen memory and learning, provided the retrieval is active on the student's behalf (Dunlosky et al., 2013). We suggest that the design of the intervention should be developed to, in the second lesson, include a short activity to retrieve content from the first lesson before introducing the tasks aiming at learning new content. Thus, we propose the proposition for RP will be developed as: *If students are supposed to use prior knowledge when learning new mathematical content, students should, in a separate lesson before learning new mathematics, have opportunities to retrieve adequate prior knowledge, and in the following lessons, start with an activity engaging them in retrieving the content from the previous lesson*.

Acknowledgement

The research is funded by Skolforskningsinstitutet.

References

Cobb, P., Jackson, K. & Dunlap, C. (2016). Design research: an analysis and critique. In L. D. English & D. Kirshner (Eds.), *Handbook of International Research in Mathematics Education* (pp. 481–503). Routledge.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J. & Willingham, D. (2013). Improving students' learning with effective learning techniques: promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, 14, 4–58.

Jonsson, B., Norqvist, M., Liljekvist, Y. & Lithner, J. (2014). Learning mathematics through algorithmic and creative reasoning. *Journal of Mathematical Behavior*, 36, 20–32. doi: 10.1016/j.jmathb.2014.08.003

Jonsson, B., Wiklund-Hörnqvist, C., Stenlund, T., Andersson, M. & Nyberg, L. (2020). A learning method for all: the testing effect is independent of cognitive ability. *Journal of Educational Psychology*. doi: 10.1037/edu0000627

Kirschner, P. A., Sweller, J. & Clark, R. E. (2006). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41 (2), 75–86. doi: 10.1207/s15326985ep4102_1

Lester, J. & Cai, F. K. (2016). Can mathematical problem solving be taught? Preliminary answers from 30 years of research. I P. Felmer et al. (Eds.), *Posing and solving mathematical problems: research in mathematics education* (pp. 117–135). Springer.

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

Norqvist, M., Jonsson, B., Lithner, J., Qwillbard, T. & Holm, L. (2019). Investigating algorithmic and creative reasoning strategies by eye tracking. *The Journal of Mathematical Behavior*, 55, 100701. doi: 10.1016/j.jmathb.2019.03.008

Olsson, J. & D'Arcy, D. (2022). Students' reasoning and feedback from a teacher. *Nordic Studies in Mathematics Education*, 27 (1), 27–49.

Olsson, J. & Granberg, C. (2019). Dynamic software, task solving with or without guidelines, and learning outcomes. *Technology, Knowledge, and Learning*, 24 (3), 419–436.

Olsson, J. & Granberg, C. (2022). Teacher-student interaction supporting students' creative mathematical reasoning during problem solving using Scratch. *Mathematical Thinking and Learning*, 1–28. doi: 10.1080/10986065.2022.2105567

Schoenfeld, A H. (2007). Method. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 69–107). Information Age.

Schoenfeld, A. H. & Sloane, A. H. (Eds.). (2016). *Mathematical thinking and problem solving*. Routledge.

Swan, M. (2008). Designing a multiple representation learning experience in secondary algebra. *Educational Designer*, 1, 1–17.

Sweller, J. (2020). Cognitive load theory and educational technology. *Educational Technology Research and Development*, 68 (1), 1–16. doi: 10.1007/s11423-019-09701-3

Investigating the role of data in the teaching of statistical modelling – a literature review

Takashi Kawakami and Jonas Bergman Ärlebäck

This study develops and discusses a framework for identifying the role of data in empirical research studies, focusing on the teaching of statistical modelling based on a literature review. A theoretically derived framework that recognises three distinct roles of data in teaching statistical modelling was used to analyse 63 papers. The findings show that data are assigned three different roles: data as a source of models, data as evidence to test model validity, and data as a storyteller of the data-generating process. The results of the analysis are further discussed in relation to three rationales for teaching statistical modelling: competency oriented, content oriented, and socially oriented. Directions for further theoretical and empirical lines of enquiry regarding the teaching of statistical modelling are suggested.

The importance and role of data are expanding in our data-driven society (e.g. Burrill & Pfannkuch, 2023; Lesh et al., 2008), demanding citizens to be productive in their personal and professional lives and to participate in society in a well-informed and critical way (Gal & Geiger, 2022). One approach to support students in coping with this data-intensive world is to develop their data literacy – the ability to describe, reason, analyse, argue and make decisions using data (OECD, 2019), which is necessary for all citizens.

Additionally, the recent rapid technological developments have diversified how data can be collected, managed and organised in large quantities and how the quality of the collected data can be improved (Burrill & Pfannkuch, 2023). Consequently, there is a need to start rethinking the concept of data and its role in mathematics and statistics education. This means, for example, considering non-traditional data types such as spatial data, video and images (Lee & Wilkerson, 2018). Another consequence is an increased emphasis on data science education (Burrill & Pfannkuch, 2023). One of the questions that arise in these evolving educational contexts is how school mathematics (including statistics education) can contribute to support students in developing the required competencies in data science and data literacy.

Takashi Kawakami, Utsunomiya University
Jonas Bergman Ärlebäck, Linköping University

Previous research in mathematics and statistics education has explored the role and use of modelling, which puts data at the core of activities in educational contexts (Ärlebäck & Kawakami, 2023; Frejd & Ärlebäck, 2021; Kawakami & Ärlebäck, 2022). Notably, this line of inquiry has, in addition, (i) broadly pointed out the importance of data in mathematical and statistical modelling activities (e.g. Blomhøj & Jensen, 2007; Pfannkuch et al., 2018); and, as elaborated below, (ii) identified three rationales for teaching statistical modelling (*competency oriented, content oriented* and *socially oriented*) (Ärlebäck & Kawakami, 2023; Kawakami & Ärlebäck, 2022). However, the role of data in developing competencies, concepts, and insights through statistical modelling with real-world and societal contexts, has not been systematically investigated and articulated. This study aimed to (i) develop and discuss a framework for identifying the roles data have been reported to play in empirical studies that investigated the teaching of statistical modelling based on a literature review; and (ii) to provide an overview of the distributions of these roles.

Theoretical framework and research question

Statistical modelling and three rationales for teaching it

No widely accepted definition of statistical modelling exists in statistical education. For the purposes of this review, we adopted the general definition of statistical modelling provided by Langrall et al. in their review of statistical education research:

> Modeling in statistics refers to any one of a number of practices: the development of a distribution (empirical or descriptive model) from data; the process of creating a theoretical (probability) model from an empirical model; and the practice of sampling from a theoretical model (simulation).
> (Langrall et al., 2017, p. 502)

Pfannkuch et al. (2018) opined that statistical modelling lies on a spectrum between being solely data-driven and being solely theory-driven. One end of this spectrum, called ”data modelling with graphs” by Pfannkuch et al. (2018, p. 115), involves making and/or using graphs and other representations (such as empirical models of the distribution) to make informal statistical inferences based on a good fit with data (Makar & Rubin, 2009). The other end of the spectrum is characterised as ”chance modelling with mathematical theoretical distributions” (Pfannkuch et al., 2018, p. 115), which involves making statistical inferences based on probability distribution (as theoretical models of the distribution) or simulations. In statistics education, various statistical modelling activities (such as data modelling, software modelling, exploring the output of pre-built models, and model recognition), are practised within the spectrum between these two ends (Pfannkuch et al., 2018).

Based on a systematic literature review of 48 peer-reviewed empirical research papers on teaching statistical modelling in the context of mathematics and statistics education, three rationales for teaching statistical modelling have been identified (Ärlebäck & Kawakami, 2023; Kawakami & Ärlebäck, 2022). These rationales for teaching statistical modelling are *competency oriented, content oriented* and *socially oriented* (table 1), and they capture different foci and emphasis with respect to the systematic design of statistical modelling tasks as well as the appropriate positioning statistical modelling practices within mathematics curricula. Statistical modelling can be comprehended holistically as a vehicle and platform for developing statistical competencies, statistical content, and social and societal decision-making.

Table 1. *The three rationales for teaching statistical modelling (Ärlebäck & Kawakami, 2023; Kawakami & Ärlebäck, 2022)*

#	Rationale	Description
R1	Competency oriented	This rationale aims to develop statistical competencies, including statistical literacy, reasoning, and thinking, as well as statistical processes such as statistical inquiry and informal statistical inference. It highlights the applicability of statistical modelling in solving real-world problems while emphasising the use of actual and authentic data to stress the applied nature of statistics.
R2	Content oriented	This rationale promotes the learning of statistical knowledge and concepts, such as variability, distribution, sample, and sampling, along with knowledge and concepts related to statistical models and modelling. Statistical modelling is viewed as an epistemic practice of statistics and a pedagogical tool.
R3	Socially oriented	This rationale promotes decision-making in the real world, social, and societal contexts where data are embedded and perform a crucial function. Furthermore, it emphasises the need to develop a critical understanding of the use and role of statistics, statistical models, and modelling in such contexts. Statistical modelling is viewed as both a means and an object of social criticism and decision-making based on data. It puts particular emphasis on addressing social issues and contexts.

The potential use of data in statistical modelling

Data are essential in statistical modelling and therefore, in order to achieve the educational goals of statistical modelling based on table 1, the use of data appropriate to the goals needs to be incorporated into the design of tasks and lessons. According to Shibata (2015, p. 8), the broad dictionary meaning of data is ”information on which inferences are based”, while the narrower meaning refers to ”a sequence of values of a variable or a sequence of pairs of values of multiple variables”. In statistics education in general, data are commonly utilised in the latter sense, encompassing qualitative data (e.g. textual information such as blood type), and quantitative data (e.g. numerical information such as height and temporal data in the form of time series data acquired along changes, for instance, daily temperatures) (Bargagliotti et al., 2020). Furthermore, real data, realistic data or simulated data are employed in statistics education.

It is important to note that data represent ”numbers with context” (Cobb & Moore, 1997, p. 801). This means that either a real-world, concrete, or theoretical context underlies the data, and it is imperative to understand and interpret data in one of these contexts when handling such information. Furthermore, data have a signal and noise structure (Konold & Pollatsek, 2002). Signal refers to regular patterns or central tendencies in the data, whereas noise refers to irregular random patterns or errors in the data. Data are subjected to mathematisation to unveil or abstract its underlying structure for solving real-world problems, accompanied by sources of statistical and mathematical knowledge and concepts (Lesh et al., 2008; Wild & Pfannkuch, 1999). Hence, statistical modelling demands ongoing transitions between context, data, and the (statistical) model, as depicted in figure 1 (Wild & Pfannkuch, 1999). Figure 1 illustrates that, on the one hand, data are generated from the context, and (statistical) models are generated from the data. On the other hand, data are interpreted and validated based on the context, and (statistical) models are interpreted and validated based on the data.

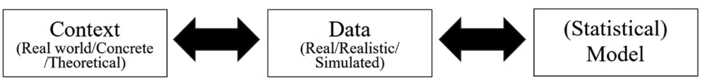

Figure 1. *Relationship between context, data, and (statistical) model (adapted from Wild and Pfannkuch, 1999)*

Data play key roles in generating, validating, and modifying models in statistical modelling or modelling in general. Dunn and Marshman (2020), identified two roles of real data in model development in educational settings: (a) to validate the model’s structure or parameter values; and (b) to estimate the model’s structure or parameter values. These two roles correspond to the arrows between the data and the (statistical) model in figure 1: data serve as *the source of the model*, and are *the evidence against which the model’s validity can be tested*. Moreover, data can also be considered as a model of phenomenon/theory (Wilkerson & Laina, 2018). The investigator does not always collect data, but often acquires them from other sources, such as public datasets. Simulated data are generated using simulation software that embeds the theory of probability distributions. As the data used in the classroom may already have been processed, ”[m]aking meaning from someone else’s data requires being aware of the measurement process that generated data” (Rubin, 2020, p. 156). Adopting such a mindset cultivates critical thinking regarding data and (statistical) models, which may result in modifications to the data collection and (statistical) models. The third role of data can be regarded as *a storyteller of the data-generating process*. This role corresponds to the arrows connecting the context and the data in figure 1. Thereby, we adopted the following three categories of potential roles of data in statistical modelling (table 2), with the aim of

Table 2. *The potential roles of data in statistical modelling*

Role	Description
Role A	Data as source of models (Data → Model)
Role B	Data as evidence by which to test model validity (Data ← Model)
Role C	Data as a storyteller of the data-generating process (Context ⇄ Data)

facilitating of the purposeful use of data according to the three rationales for teaching statistical modelling.

Research question

Given the theoretical framing of the different roles data can play in statistics education research as summarised in table 2, we formulated and addressed the following research question: *How are the roles of data (as a source, as evidence, as a storyteller) distributed in empirical research on statistical modelling that invokes different rationales for teaching statistical modelling?*

Methodology

In Kawakami and Ärlebäck (2022), we identified empirical research literature on statistical modelling in mathematics and statistics education that invokes different rationales for teaching it. The literature dataset identified from this prior study was used to conduct a systematic review of the role of data in teaching statistical modelling. The literature was selected because it demonstrated a rationale for the use of statistical modelling in mathematics and statistics education and also focused on data, as indicated in the paper selection process described below (figure 2), making it suitable for answering the research question. The literature dataset includes peer-reviewed research literature that used or investigated statistical modelling from the following influential mathematics education journals: Educational Studies in Mathematics (ESM) (1968–July 2021), ZDM: Mathematics Education (ZDM) (1997–July 2021), Mathematical Thinking and Learning (MTL) (1999–July 2021), Journal for Research in Mathematics Education (JRME) (1970–July 2021), and Journal of Mathematical Behavior (JMB) (1995–July 2021). Additionally, research papers from the following internationally recognised journals in statistics education were included: Statistics Education Research Journal (SERJ) (2002–July 2021) and Journal of Statistics Education/Journal of Statistics and Data Science Education (JSE/JSDE) (1993–July 2021). The varying start dates of the journal search is due to the fact that the journals' first issue appeared in different years. Notably, special issues on statistical modelling research were found in ZDM Vol. 50, No. 7 and SERJ Vol. 16, No. 2.

We selected eligible papers through the journals' databases, including Springer Link, JSTOR, Taylor & Francis Online, Science Direct and the IASE website. We extracted original papers that contain the terms ”data”, ”statistics”, or ”statistical”, and ”model” or ”mode[l]ling” in the article’s title, keywords or abstract. Upon closer inspection, we excluded papers that contained these terms but did not focus on empirical research. We also excluded papers focusing solely on probability modelling or on modelling learning and cognition to ensure a focused discussion. A total of 63 papers were extracted, including 7 ESMs, 16 ZDMs, 10 MTLs, one JRME, one JMB, 14 SERJs and 14 JSEs/JSDEs [1]. Figure 2 illustrates the paper selection.

For the systematic review, we read all 63 identified papers and categorized the intended function or role of data in the teaching statistical modelling in these papers by coding them into one or more of the three categories designated as roles A, B, and C in table 2. The coding criteria were based on the explicit description(s) of the function or data in the papers. We focused on (1) the description of the study’s purpose and position, (2) the intentions and purposes of the utilised teaching materials, curriculum, and teaching practices, as well as (3) the research questions presented in the papers. Table 3 presents examples of such descriptions and their coding. The first author conducted the first analysis, which included classification, and the second author independently checked the assigned papers and the analysis. Where discrepancies occurred, the authors discussed and resolved them. In Kawakami and Ärlebäck (2022), we identified the rationales in table 1 for using statistical modelling in the literature dataset, but did not analyse the roles of data in empirical research on statistical modelling that invokes different rationales for teaching statistical modelling. To answer the research question, the current study cross-tabulated the roles of data with the identified rationales in the literature dataset.

Results

We present the results of the literature review in the research question, organised in two sections: *the role of data discerned in empirical research on*

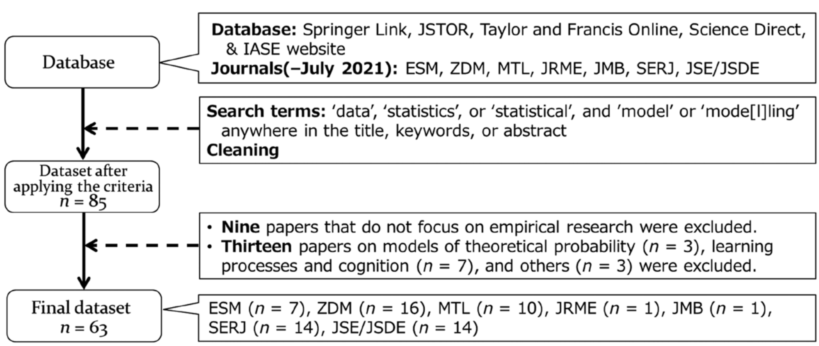

Figure 2. *Paper selection process (adapted from Kawakami and Ärlebäck, 2022, p. 1112)*

Table 3. *Examples of the descriptions for determining categories roles A, B, and C*

Category	Examples of the descriptions in the papers
Role A	– A pedagogical modelling approach that views real data as a source for a model of a situation in the real world can serve as a bridge between data and probability … (Aridor & Ben-Zvi, 2017, p. 41) – For statistical modelling, data from real world systems are a key part of building models, where the purpose is to create models that mimic random behaviour in real world systems … (Patel & Pfannkuch, 2018, p. 1198)
Role B	– The effectiveness of this model in producing the required output is evaluated by using graphs to compare the generated data with the initial data that the pupils collected or explored, in the style of AG [Active Graphing]. (Ainley & Pratt, 2017, p. 20) – Model evaluation drives model revision, which generally increases the explanatory power of a model and the scope of application as the model is fine-tuned to account for new data (i.e. evidence in scientific inquiry). (Doerr et al., 2017, p. 89)
Role C	– Typically, these stories are conceptualized as embedded within data, to be discovered or unlocked by learners. (Wilkerson & Laina, 2018, p. 1224) – In particular, the inferential explanations were grounded in the knowledge of context that was used as a story behind the data as well as enabling students to go beyond the data to make inferences. (Kazak et al., 2021/2023, p. 30)

statistical modelling, and *the role of data in terms of different rationales for the teaching of statistical modelling.*

Role of data discerned in empirical research on statistical modelling

Figure 3 summarises the ways data were used in the 63 papers examined, with the majority (45 %) of the papers employing data in tandem with both roles A and B simultaneously. The second largest proportion of papers employed data in tandem with all three roles (24 %). When examining the frequency with which data were used according to the three individual roles in the papers, it transpired that role A was the most common (97 %), followed by role B (71 %) and role C (41 %).

Upon scrutinising the 26 papers using data in line with role C (data as a storyteller of the data-generating process, Context ⇄ Data) more closely, we found

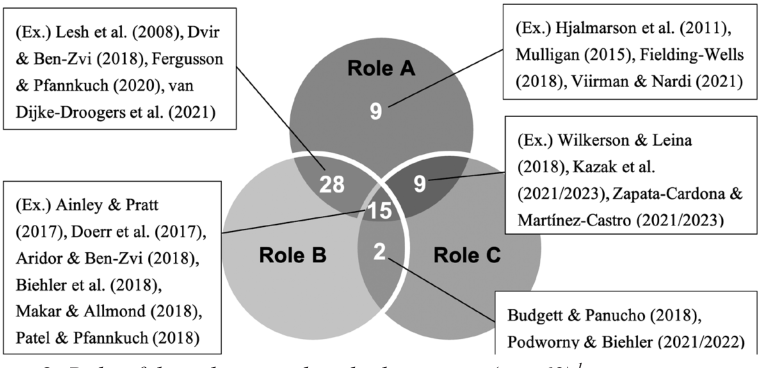

Figure 3. *Role of data discerned in the literature (n = 63)* [1]

a split between studies that focused on transitions between real-world contexts and data (*n* = 19), and studies that emphasised transitions between theoretical contexts (the world of probability distributions) and data (*n* = 7).

The former group of papers (*n* = 19) typically focused on data modelling, creating empirical or descriptive models of distributions from real or realistic data, or were included as part of the overall activity. However, details of the intended role of data in these papers were varied: raising awareness of uncertainty (*n* = 2); critical examination of data (*n* = 4); making sense of the data (*n* = 3); facilitating model development (*n* = 1); understanding of problem situations (*n* = 8); and understanding the causes of variation (*n* = 1).

The latter group of papers (*n* = 7) aimed to use simulation software to generate simulated data from a model that resembled the behaviour of the population used in the examples in the papers, with a focus on modelling chance, where the role of the data was to facilitate learners' understanding of the process by which a sample is generated from the population.

Role of data in terms of different rationales for the teaching of statistical modelling

Table 4 presents a cross-tabulation of the 63 papers analysed, showing how the rationales for teaching statistical modelling identified by Kawakami and Ärlebäck (2022) relate to the roles of data identified in the papers. Rationale R1 (competency oriented) were used by 87 % of the papers that adapted all three of roles A, B and C. For papers that used rationale R2 (content oriented), all but two (95 %) used data in line with role A. Half of the 10 papers that used rationale R3 (socially oriented) adopted role C.

Table 4. *Cross-tabulation of the role of data and three rationales for teaching statistical modelling discerned in the literature (n = 63)*

	Role A	Roles A and B	Roles A and C	Roles B and C	All roles	Total
R1	3	11	2		3	19
R2	2	9			2	13
R1 and R2	3	5	2	2	9	21
R1 and R3		1	3		1	5
R2 and R3	1		1			2
R1, R2, and R3		2	1			3
Total	9	28	9	2	15	63

Discussion and conclusion

The result presented in figure 3, namely that the most common use of data in the papers analysed was according to roles A and B (data as a source for models and as evidence to test model validity), corresponds with the observation made by

Cobb and Moore (1997, p. 810) that ”[s]tatistics in practice resembles a dialogue between models and data”. Moreover, regarding the second-most common use of data in the papers analysed, all three roles were in line with the conclusion by Pfannkuch et al. (2018, p. 1116), which recognised data as central in the teaching of statistical modelling: ”[C]ontexts interact and play an important role in promoting students' learning and reasoning from data”.

Table 4 reflects that ore than half of the papers in which R1 (competency oriented) was the only rationale identified (n = 19), included roles A and/or B. Data used in accor-dance with role C featured in only five papers (26 %), and then never as the only role. These results are also in line with Cobb and Moore's (1997, p. 810) finding: ”Statistics in practice resembles a dialogue between models and data”. If one of the aims of mathematical modelling is the competency to move between the real world and the mathematical world (e.g. Blomhøj & Jensen, 2007), then one of the aims of statistical modelling may be the competencyto move between the data world and the model world (Pfannkuch et al., 2018).

The fact that almost all the papers in which R2 (content oriented) was the only rationale for teaching statistical modelling had adopted role A, illustrates the importance of the learner's confidence in making models from the data when constructing statistical contents.

In the socially oriented papers (R3), using data in line with role C, the focus was on making sense of and critically examining data and the context behind the data. With the growth of big and social data, critical examination of the processes by which real data are generated (e.g. Wilkerson & Laina, 2018) will become increasingly important.

This paper presented a framework for identifying the role of data in empirical research studies, focusing on the teaching of statistical modelling based on a literature review. The framework provides a perspective to facilitate researchers and teachers to identify the purposeful use of data according to the rationale for teaching statistical modelling and to compare the positions and characteristics of existing statistical modelling practices. Given the diverse ways in which high-quality data can be collected, managed and organised in large quantities (Burrill & Pfannkuch, 2023), and the need to rethink the concept of data and its role in mathematics and statistics education (Lee & Wilkerson, 2018), this analysis could be extended by conducting a more qualitative study of the role and types of data actually used in teaching statistical modelling. This would provide a more nuanced picture and suggest directions for further theoretical and empirical lines of enquiry regarding the teaching of statistical modelling.

References

Ainley, J. & Pratt, D. (2017). Computational modeling and children's expressions of signal and noise. *Statistics Education Research Jou*rnal, 16 (2), 15–37. doi: 10.52041/serj.v16i2.183

Aridor, K. & Ben-Zvi, D. (2017). The co-emergence of aggregate and modelling reasoning. *Statistics Education Research Journal*, 16 (2), 38–63. doi: 10.52041/serj.v16i2.184

Ärlebäck, J. B. & Kawakami, T. (2023). The relationship between statistics, statistical modelling and mathematical modelling. In G. Greefrath et al. (Eds.), *Advancing and consolidating mathematical modelling: research from ICME-14* (pp. 293–309). Springer.

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S. et al. (2020). *Pre-K-12 guidelines for assessment and instruction in statistics education II (GAISEII)*. American Statistical Association.

Blomhøj, M. & Jensen, T. H. (2007). What's all the fuss about competencies? In W. Blum, P. Galbraith, H.-W. Henn & M. Niss (Eds.), *Modelling and applications in mathematics education. The 14th ICMI study* (pp. 45–56). Springer.

Burrill, G. & Pfannkuch, M. (2023). Emerging trends in statistics education. *ZDM*, 56, 19–29. doi: 10.1007/s11858-023-01501-7

Cobb, G. W. & Moore, D. S. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly*, 104, 801–823. doi: 10.2307/2975286

Doerr, H. M., delMas, R. & Makar, K. (2017). A modeling approach to the development of students' informal inferential reasoning. *Statistics Education Research Journal*, 16 (2), 86–115. doi: 10.52041/serj.v16i2.186

Dunn, P. K. & Marshman, M. F. (2020). Teaching mathematical modelling: a framework to support teachers' choice of resources. *Teaching Mathematics and its Application*, 39, 127–144. doi: 10.1093/teamat/hrz008

Frejd, P. & Ärlebäck, J. B. (2021). Connections and synergies between the teaching and learning of statistics and modelling: a pilot study. In F. K. S. Leung et al. (Eds.), *Mathematical modelling education in east and west* (pp. 607–617). Springer.

Gal, I. & Geiger, V. (2022). Welcome to the era of vague news: a study of the demands of statistical and mathematical products in the COVID-19 pandemic media. *Educational Studies in Mathematics,* 111 (1), 5–28. doi: 10.1007/s10649-022-10151-7

Kawakami, T. & Ärlebäck, J. B. (2022). Characterising rationales for using statistical modelling in education research from a mathematical modelling perspective: a review. In J. Hodgen et al. (Eds.), *Proceedings of CERME12* (pp. 1109–1116). Free University of Bozen-Bolzano and ERME. https://hal.science/hal-03759036

Kazak, S., Fujita, T. & Turmo, M. P. (2021/2023). Students' informal statistical inferences through data modeling with a large multivariate dataset. *Mathematical Thinking and Learning*, 25 (1), 23–43. doi: 10.1080/10986065.2021.1922857

Konold, C. & Pollatsek, A. (2002). Data analysis as the search for signal in noisy processes. *Journal for Research in Mathematics Education*, 33 (4), 259–289. doi: 10.2307/749741

Langrall, C., Makar, K., Nilsson, P. & Shaughnessy, J. M. (2017). Teaching and learning probability and statistics: an integrated perspective. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 490–525). NCTM.

Lee, V. R. & Wilkerson, M. (2018). *Data use by middle and secondary students in the digital age: a status report and future prospects*. Commissioned Paper for the National Academies of Sciences, Engineering, and Medicine, Board on Science Education, Committee on Science Investigations and Engineering Design for Grades 6-12.

Lesh, R., Middleton, J. A., Caylor, E. & Gupta, S. (2008). A science need: designing tasks to engage students in modeling complex data. *Educational Studies in Mathematics*, 68, 113–130. doi: 10.1007/s10649-008-9118-4

Makar, K. & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal*, 8, 82–105. doi: 10.52041/serj.v8i1.457

OECD (2019). *OECD future of education and skills 2030 concept note*. https://www.oecd.org/education/2030-project/contact/

Patel, A. & Pfannkuch, M. (2018). Developing a statistical modeling framework to characterize Year 7 students' reasoning. *ZDM*, 50 (7), 1197–1212. doi: 10.1007/s11858-018-0960-2

Pfannkuch, M., Ben-Zvi, D. & Budgett, S. (2018). Innovation in statistical modelling to connect data, chance and context. *ZDM*, 50 (7), 1113–1123. doi: 10.1007/s11858-018-0989-2

Rubin, A. (2020). Learning to reason with data: How did we get here and what do we know? *Journal of the Learning Sciences*, 29, 154–164. doi: 10.1080/10508406.2019.1705665

Shibata, R. (2015). *Data assay and data science* (in Japanese). Kindai kagaku sha.

Wild, C. J. & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67 (1), 223–265. doi: 10.1111/j.1751-5823.1999.tb00442.x

Wilkerson, M. H. & Laina, V. (2018). Middle school students' reasoning about data and context through storytelling with repurposed local data. *ZDM*, 50 (7), 1223–1235. doi: 10.1007/s11858-018-0974-9

Note

1 For a list of the full references of all 63 papers in the literature data set, see https://drive.google.com/file/d/1H1eLQ6OGJSq3XQE5pmFto5PSExCF2uq_/view?usp=sharing

Communication strategies when reading less familiar mathematical expressions aloud

Ulrika Wikström Hultdin

Spoken mathematical symbols are part of oral communication in mathematics classrooms. The transition from written mathematical symbols into spoken sound is an intricate process that, in some respects, parallels the translation of one language into another. This study focuses on strategy-use in the reading aloud of mathematical symbols and compare them to strategies found in second language communication (SLC). In an analysis of sound recordings of university students who read symbolic expressions aloud, SLC strategies were found to be abundant. This indicates that some obstacles encountered when transforming written symbols into spoken words are similar to problems occurring in SLC. It is suggested that mathematics teaching should include some instructions on how to read symbolic expressions.

Some say that mathematics is a language, and others think that is not the case (Wakefield, 2000). Nevertheless, in mathematics education research, language-rich classroom activities are generally acknowledged as central to learning (Morgan et al., 2014). More specific suggestions on the topic have included teaching mathematics using methods common to second language learning (Wakefield, 2000; Bossé et al, 2018; Ledibane et al, 2018). Because both language use and the use of mathematics are considered social human activities that are learned in communication with others, this is in tune with other current ideas, for example, about the benefits of collaborative learning (Laal & Ghodsi, 2012).

Envisioning the learning of mathematics in a functional collaborative environment typically includes students working together on mathematical problems, discussing and reasoning about concepts and methods, sharing ideas, interpreting different visuals, and so on. It might also include students' advancements in the use and understanding of mathematical symbols. In such a collaborative environment, the students will probably, now and then, need to transform written mathematical expressions into speech. Intriguingly, little is known about what happens in such situations, and due to the characteristics of the mathematical symbol system, it might involve some challenges.

Ulrika Wikström Hultdin
Umeå University

Mathematical symbols are used to represent values, operations, functions, and patterns. It is a written system, where the signs relate to their content by convention (Pimm, 1987). In contrast to the reading of alphabetical languages, where words can be spelled out syllable by syllable, the reading aloud of mathematical symbols requires the reader to have experience either with the sound, or an alphabetical translation, of the symbols before they can read themselves. An additional peculiarity is that a reader, despite knowing the name of every single symbol included in a mathematical expression, can still be unable to perform the interpretation necessary to express it orally. Differences in the size and spatial location of a symbol can completely change its meaning and how it is read. Consequently, the reading becomes different from the reading of the ordinary Swedish language. For example, the reading order is not necessarily left-to-right and top-to-bottom. Furthermore, there is often more than one option on how to write an expression. Compare, for example, $\frac{x^2}{5}$, $x^2/5$ and (x^2)/5. The possibility to read an expression in several ways is also common. Consider, for example, y'' which can be read as ”y-double-prime” just as well as ”the second derivative of y”. Both are mathematically acceptable. Yet, the latter alternative is a more interpretive form of reading than the former (Pimm, 1987).

At the time I started to develop an interest in reading in mathematics, previous research had shown that silent reading of texts with mathematical symbols interwoven in sentences was more difficult than the reading of texts without any mathematical symbols (Österholm, 2006). Intrigued by those results, I did some work focused on the variation in word use in reading aloud of mathematical symbols to find hints on what could be disturbing the readers (Hultdin, 2013). While listening to some recordings of students that were reading mathematical expressions, I surprisingly found their struggle to be very familiar; were they not behaving in a similar way to persons trying to express something in a second language? Considering the special characteristics of the symbol system, it is not a farfetched assumption to think that students will sometimes experience a lack of vocabulary, or challenges with the interpretation, when transforming a somewhat unfamiliar written mathematical expression into speech. This could cause a stall, which has to be overcome, and to overcome it, some strategy is needed.

Second language communication strategies

To experience a lack of vocabulary or grammar difficulties that threaten to interrupt or impair communication is common for second language (L2) speakers. In such situations, a plethora of second language communication (SLC) strategies have been found to come into service. Different taxonomies describe different strategy types under different names but, in the end, there seem to be more similarities than differences between these descriptions (Avval, 2012). In 1997, Dörnyei and Scott summarized all the communication strategies proposed to date and divided them into three main groups: direct strategies, interactional strategies, and indirect strategies. Direct strategies concern, for example, change in vocabulary but also self-repair and mumbling. Interactional strategies include asking for help, clarification, and repetition. The indirect strategies are performance-related or related to time pressure and include, for example, marking uncertainty about the message (performance-related), or using filling words to keep the communication channel open (time pressure-related).

Not only in situations of oral communication but also when written language is translated from a first language to L2, SLC strategies come into play (Rabab'ah, 2008). Transforming symbolic expressions into speech bears some resemblance to the situation when one written language is translated to another, at least on the surface. First, the starting point is a written text. Second, the written symbols, and sometimes even the meaning of them, can be more familiar than their ”translation”. Third, there is not always a one-to-one transformation; a symbolic expression can often be transformed into several adequate oral expressions.

In the present study, the purpose is to gain knowledge about what happens in a situation where students read out symbolic expressions with which they are not completely familiar. The focus is on the transformation of written symbols into speech, and the question addressed is: What strategies can be identified when students read less familiar mathematical expressions aloud? The aim is to illustrate some of the complexity in the transformation process, not to provide an exhaustive enumeration of all strategies possible to find when students struggle with the reading of symbols. The results could propose ideas for strategies to overcome the struggle.

Methods

Although reading aloud does not cover the whole spectrum of classroom situations in which students transform mathematical expressions into speech, it allows for the investigation of many diverse expressions in a short time. Since the focus of the study was on the transformation and not on communication in general, the data collection could be limited to sound recordings. All data was originally collected as part of a more extensive project thesis (Hultdin, 2013) [1]. All participants gave consent for the use of the recorded material in research.

The participants were 18 students (male = 14, female = 4, age: 19–25 years), all fluent Swedish speakers, taking a preparatory university course in mathematics. The texts that were read were short Swedish texts, 3–8 lines. Each participant was recorded in a separate session. During the sessions, which were taking place in an office near the students' ordinary classroom, only the participant and the instructor were present. The data was collected one week before the final course exam. In a break between lessons, potential study participants were informed about the study focus, time and place for participation, confidentiality, and the general procedure of the reading session. The texts read were modifications of ordinary textbook paragraphs or task texts and included mathematics taught in the current course. Of 12 texts read, the six causing the students the most problems (a–c below) were selected for analysis. Three integral expressions, two differential equations, a primitive function, and double angle identities were included in the selection.

At the time, the students were divided into three teaching groups, taught by three course teachers. The analysis included a comparison of participants' readings to *conventional* reading. For this purpose, a standard was created. By asking the course teachers to write down how they would read a few selected mathematical expressions and specifically asking them to come up with as many alternatives as possible that they would judge to be correct, various ”standard readings” were collected.

It was assumed that students would use various communication strategies when struggling with their reading of mathematical symbols. Therefore, strategy use was analyzed when one of the following occurred in the sound recordings: a) longer pauses (compared to reading other text parts), b) use of non-standard vocabulary, or c) a non-standard reading order (compared to the ”standard readings”).

Analysis based on Dörnyei and Scott's taxonomy

Because of the assumed similarities between the reading situation and L2 communication, the data was classified with the help of Dörnyei and Scott's taxonomy of SLC strategies (1997). It comprises a summary of relevant SLC strategies, and has a focus on produced verbal interaction, which made it suitable for this first exploration of strategy use for readers of mathematical texts. The choice foregrounds similarities that can be found between the reading situation and a situation of L2 communication. However, also non-SLC strategies were noted during the analysis.

Since the taxonomy concerned oral L2 communication involving a (more active) second person, all the interactional strategies such as ”other repair”, ”other repetition” and ”feigning understanding” had to be excluded. The type of data collected also limited the number of strategies possible to identify in the material. Because all participants were Swedish speaking, strategies used to overcome problems voicing specific L2 sounds were excluded. The strategy where ”more words were used to achieve a particular communicative goal than what is considered normal in similar L1 situations” (termed ”over explicitness” in Dörnyei and Scott, 1997), turned out to be either literal translations or paraphrasing when symbols were read aloud. Thus, it was not included as a separate category. The final 18 categories are described below.

Direct, resource-deficit-related

Message abandonment. Giving up and skipping the expression/text.

Message reduction. Not reading single words or longer text chunks in the middle of a text.

Message replacement. When saying anything (instead of nothing), but not at all with the intended meaning. An example would be to say ”and-lah-lah-lah” instead of the symbols.

Non-standard readings were classified as one of the three above only when no other strategy, for example, ”omission”, could be assigned.

Restructuring. Abandoning the first communication plan. Continuing with the same message but read in another way. Example from reading y'': ”[...] y-prime, y ... let’s see, y ... the derivative twice”.

Approximation. Using a superordinate or relative term as an alternative to a missing one. For example, saying ”area” instead of ”integral”.

Literal translation. Using a known pronunciation of a symbol in a context where it is normally not used. For example, reading out all parentheses.

Paraphrasing. Reading in a more explanatory way after stalling or hesitation. For example, saying ”the primitive function of f-of-x” when reading $F(x)$. This strategy can include both standard and non-standard vocabulary.

Word coinage. Unconventional reading of a mathematical expression based on previous experiences of mathematical syntax and vocabulary. Example: The division sign in $\frac{dx}{dy}$ in an integral expression is read ”divided by”.

Use of similar-sounding words. Substituting an unknown word (either existing or non-existing) with one that sounds, more or less, like it.

Use of all-purpose words. Using all-purpose words, such as ”this thing” or ”whatever it’s called” instead of specific ones.

Mumbling. Swallowing or muttering a word inaudibly.

Omission. Skipping a symbol and continuing as if it had been read.

Retrieval. When retrieving a lexical item, saying a series of incomplete or wrong forms before reaching the optimal one.

Direct, performance-problem-related

Self-rephrasing. Repeating a term and, at the same time, changing it (without changing the message). Example: ”the second derivative ... the second derivative of f-of-x”

Self-repair. Self-correcting, either directly or after repetition of the incorrect word/phrase.

Self-repetition. Repeating a word or phrase immediately after it is said.

Indirect, time-pressure related

Use of fillers. Using all-purpose words, or sounds, to fill pauses.

Indirect, performance-problem-related

Verbal strategy marker. Using verbal marking phrases, such as ”I don’t know what to say.”, before or after a strategy to signal that the word or phrase might not carry the intended meaning.

For quite a few of the readings, follow-up interviews would have been necessary to report the exact strategy in use. For example, if a participant says the word ”function” instead of ”integral” it is not possible to know if the person sees the concepts as similar and, thus, uses an approximation, or if they consciously replace the original message just to say something. Yet, many readings were to the point and could easily be classified based on the descriptions. Because of the limitations to the method, only the occurrence – if a strategy was present in data or not – was reported for each strategy, and not its frequency.

Results

When transforming written mathematical symbols into speech, numerous strategies were identified. Strategy use was often associated with the difficulties encountered with specific expressions. When reading integral expressions, about half of the participants exchanged the integration limits. When learning to read in Swedish or English, we are trained to read top-to-bottom. This reading order was also used for symbolic expressions. However, the conventional reading order for integral limits is the opposite, bottom-to-top. The strategy identified was not so much a strategy as a preservation of reading order, and certainly not one of the listed SLC-strategies that were based on oral communication.

Another non-SLC strategy was found for the reading of the symbols α and β, which were symbols unfamiliar to many of the participants. Instead of ”alpha”, α was often read like the letter ”a”, and in one case, ”ex” (x). Beta was read like the letter ”b”. The strategy used in these specific cases was to substitute the unknown pronunciation with the known pronunciation of a symbol of a similar shape. For natural languages where the writing is based on alphabetic sounds, the strategy when reading an unknown written word could instead be to spell it out letter by letter. When a symbol is not connected to a specific language sound, the strategy seems to be to focus on the shape of the symbol and approximate it with a similarly shaped one.

A majority of the strategies identified were similar to SLC strategies. Of the 18 potential strategies listed, 15 were found in the recorded material (table 1). This included both direct strategies that were related to resource deficits and

Table 1. *SLC strategies in aloud reading of unfamiliar symbolic expressions*

	Strategy type	Deviation from standard	Found (Y/N)
Direct strategies			
Resource-deficit related	Message abandonment	message/content	Y
	Message reduction	message/content	N
	Message replacement	message/content	Y
	Restructuring	vocabulary use	Y
	Approximation	vocabulary use	Y
	Literal translation	vocabulary use	Y
	Paraphrasing	vocabulary use	Y
	Word coinage	vocabulary use	Y
	Use of similar sounding word	vocabulary use	N
	Use of all-purpose word	vocabulary use	N
	Mumbling	vocabulary use	Y
	Omission	vocabulary use	Y
	Retrieval	self-correction/ vocabulary search	Y
Performance problem-related	Self-repair	self-correction	Y
	Self-rephrasing	self-correction/ vocabulary search	Y
Indirect strategies			
Time pressure related	Self-repetition	vocabulary search	Y
	Use of fillers	vocabulary search	Y
Performance problem-related	Verbal strategy marker	marking uncertainty	Y

problems with the own performance; as well as indirect strategies related either to one's own performance or to time pressure.

Strategies transforming the message in such a way that it would be impossible to interpret in the same way as the original print – *message abandonment, message reduction* and *message replacement* – were not very common. *Message abandonment* was almost seen in one case where a participant, who did not find the words, wanted to skip to the following text but did continue after some encouragement from the instructor. The *message reduction* strategy was never found to be the primary strategy. In a few cases, the message was replaced not only by an approximation but by something substantially different. In those cases, it was a matter of saying anything at all, like ”blah-bla-bla”, just to fulfil the task.

Lack of vocabulary led to the use of several strategies. The resulting message was not unrecognizable but often included some ambiguities. For example, when the word ”integral” was *approximated* with ”function”. For unconfident readers, the *mumbling* strategy, or the complete *omission* of single symbols, were common. For example, when one of the participants used the term ”top value” (Swe: ”toppvärde”) when reading the integration limits of an integral expression, that word was read less clearly and in a significantly lower voice compared to other parts of the expression. Examples of the omission strategy were found when a participant consistently omitted the integral sign in the reading of integral expressions and when others did not read the ”*dx*” at the end of the expressions.

Word coinage was a less common, but still present, strategy. One example where reading was based on previous experiences with mathematical syntax and vocabulary was when $\int_a^b f(x)\,dx$ was read as ”the integral *times f*-of-*x* dee-*x* from *a* to *b*” (Swe: ”integralen gånger f-av-x de-x från a till b”). The word choice was assumed to be based on previous experiences with implicit multiplication signs.

The unconventional reading showed a large span in the degree of interpretation of the expressions; both *paraphrasing* and *literal translation* were used. Specific to this study was the participants’ search for words when encountering y''. When this, after hesitation, was read as ”the second derivative” or ”the derivative twice”, it was interpreted as *paraphrasing*. Reading $\frac{dx}{dy}$ as ”dee-why divided by dee-tee” (Swe: ”de-y genom de-te”) was interpreted as a *literal translation*.

It was common to find simultaneous usage of multiple SLC strategies. For example, the *paraphrasing* strategy was more than once part of a *restructuring* strategy. *Restructuring* was also seen together with *self-repetition* (”from, ... from, from”), *self-rephrasing* (”limits, ... interval”), and the *use of fillers* (”ehm ... ehm ...”), strategies used to keep communication going.

In some cases, it was not completely clear which strategy was used. For example, the reading ”The equation *y*-prime-two, no, the derivat ... the second derivative of *y*” (Swe: ”Ekvationen y-prim-två, näe, derivat ... andraderivatan av y”) can be seen both as a *self-correction* indicated by the ”no” followed by a change of the word, and/or as a *retrieval* where the final word is found after saying incomplete and wrong words. Furthermore, *strategy markers*, where participants communicate their uncertainty about message content, often had the additional function of being fillers. Examples were phrases such as ”What do you say?” and ”What’s the name of this?”, said before reading a difficult passage. An example of a strategy marker that was not used as a filler was the question ”If that’ll work?” (Swe: ”Om det går bra?”) said after reading an expression.

The conclusion is that manifold strategies were incited when students read less familiar symbolic expressions aloud, and many of the strategies were similar to strategies used in second language communication.

Discussion

Seventeen strategies were identified in the recordings of students who were reading symbolic expressions aloud. Two of them were not second language communication (SLC) strategies, and they were both seen in connection with certain symbolic expressions. For integral limits, where the reading order is reversed compared to written language in general, it became obvious that preserving the top-to-bottom reading order was a strategy used when reading symbolic expressions. For singular symbols with unknown names, the strategy was to substitute them with the names of similarly shaped symbols. While top-to-bottom reading might be used without the experience of any communication problems, in specific situations, it is a strategy related to a resource deficit. That is, not knowing the reading order. The symbol shape approximation is more clearly a strategy for overcoming a difficult situation in line with a situation of SLC where vocabulary is lacking. The proper name of a symbol is not known, and that problem needs to be solved. However, the strategy used when solving the problem is not an SLC strategy.

Strategies similar to SLC strategies were abundant in the recorded material. Of the 18 potential strategies listed, 15 were found at least once. The potential SLC strategies that were not found in the data, might be found in situations where a more active communication partner is involved. That may also be the case for many of the strategies that were excluded from the study for methodological reasons.

That the strategic use of similar-sounding words did not occur in this study might be explained by the expressions included in the reading tasks. If the students did not experience any problems finding a word where they already knew a similar sounding one, the strategy could not be used. The limited number of participants is also a factor to consider in relation to the non-occurring SLC strategies. For example, a person who is more prone to using all-purpose words such as ”this thing” and ”that stuff” in everyday language might also have used them in the reading of mathematical expressions. It is not known whether the 18 participants in this study included such a person, and if so, the reading situation may not have stimulated this type of strategy. That message reduction did not occur as a strategy can be a result of the analysis method, at least partly. In many cases, the resulting message was reduced but it was also possible to find that another, often more specific, strategy was used in the situation.

According to the taxonomy of Dörnyei and Scott (1997), the occurring SLC strategies were related to resource deficits and performance problems, as well as to the experience of time pressure in the communication situation. Although the data was limited to the reading aloud of selected expressions, the results so far imply that the experience of the own performance and the time pressure in the reading situation are similar to what is experienced by L2 speakers in SLC situations. Yet, the resource deficits seem to comprise some additional categories for transformations from symbols to speech. That is not surprising because of the difference between the sign systems. For example, unknown written words can still be read aloud, and for the written Swedish language, the reading order is fixed. Nevertheless, similarities in the use of communication strategies may reflect similarities in the obstacles encountered when reading less familiar symbolic expressions and speaking in L2. Use of similar strategies could also indicate similarities between reading symbols aloud and translating from a first language to an L2, another situation where SLC strategies are activated (Rabab'ah, 2008).

For a teacher to adequately help students learn the language of mathematics – the communication in words about mathematical ideas – it is important to know the many obstacles lined up along the way. As shown, one of the difficulties that causes the need for a communication strategy when transforming written symbols into speech is insufficient vocabulary. Many teachers are already explicitly teaching vocabulary to their math students by explaining and relating different words (Riccomini et al, 2015). If more teachers became aware of the problems associated with the transformation of mathematical expressions into speech, it would not be a large step to include explicit instructions on how to read different symbolic expressions related to the topic currently taught. It is worth noting that some of the unconventional paraphrasing examples identified in the data showed that students can have a sense of the meaning of an expression, or at least its functional use, before knowing a common way to read it. One example would be when the participants are saying ”the derivative twice” (see Hultdin, 2013, for more examples).

A final reflection is that human communication is rarely bound to one semiotic system or one type of situation. We communicate by any means available, using the signs we know and the strategies we have acquired, and a common strategy in second language communication could just as well be used when reading a mathematical expression aloud. The aim above all is to deliver the message.

References

Avval, S. F. (2012). Communication strategies in translation: a review on the taxonomies from 1977 to 2011. *Journal of Siberian Federal University. Humanities & Social Sciences,* 5 (6), 768–779.

Bossé, M. J., Ringler, M., Bayaga, A., Fountain, C. & Young, E. S. (2018). Acquiring math: connecting math learning and second language acquisition. *International Journal for Mathematics Teaching and Learning*, 19 (2), 223–252.

Dörnyei, Z. & Scott, M. L. (1997). Communication strategies in a second language: definitions and taxonomies. *Language learning*, 47 (1), 173–210.

Hultdin, U. W. (2013). *Vad döljer sig bakom orden? Strategi och variation vid utläsning av matematiska symboluttryck* (Project thesis). Umeå University.

Laal, M. & Ghodsi, S. M. (2012). Benefits of collaborative learning. *Procedia-social and behavioral sciences*, 31, 486–490.

Ledibane, M., Kaiser, K. & Van der Walt, M. (2018). Acquiring mathematics as a second language: a theoretical model to illustrate similarities in the acquisition of English as a second language and mathematics. *Pythagoras*, 39 (1), 1–12.

Morgan, C., Craig, T., Schütte, M. & Wagner, D. (2014). Language and communication in mathematics education: an overview of research in the field. *ZDM*, 46 (6), 843–853.

Pimm, D. (1987). *Speaking mathematically: communication in mathematics classrooms*. Routledge & Kegan Paul.

Rabab'ah, G. A. (2008). Communication strategies in translation. *Babel*, 54 (2), 97–109.

Riccomini, P. J., Smith, G. W., Hughes, E. M. & Fries, K. M. (2015). The language of mathematics: the importance of teaching and learning mathematical vocabulary. *Reading & Writing Quarterly*, 31 (3), 235–252.

Wakefield, D. V. (2000, September). Math as a second language. *The Educational Forum*, 64 (3), 272–279.

Österholm, M. (2006). Characterizing reading comprehension of mathematical texts. *Educational studies in mathematics*, 63, 325–346.

Note

1 The report is based on part of a project thesis (Hultdin, 2013). At the time, part of the results was reported in Swedish. Examples from data in the present report are the same as in the thesis but translated from Swedish to English. All other text is newly written or further processed (methods, results) for this report and not a mere translation.

Toddlers learning the meaning of counting words as increase in discerned aspects of number

Camilla Björklund and Hanna Palmér

The focus of this paper is on toddlers (1–3-year-olds) learning the meaning of counting words while participating in a three semester preschool intervention. This was investigated with task-based interviews analysed through the lens of variation theory of learning which implies observing the aspects of numbers discerned by the toddlers. In general, the toddlers learned to use counting words purposively. Discerning representations as an aspect of counting words was shown to be central, but cardinal and ordinal meaning of counting words liberate a more advanced understanding of representations. Further, the study shows that the development of toddlers' knowledge is diverse as there are quite different learning trajectories among the toddlers in the study.

Observations of young children's engagement in preschool activities was the starting point for implementing designed numeracy education with 27 toddlers (1–3 years) during three semesters in Swedish preschools. The goal of the intervention was to make possible for toddlers to explore the meaning of numbers, particularly ordinality, cardinality, part-whole relations and representations, as these have been found to be necessary to discern in order to develop a powerful basic number knowledge (Björklund et al., 2021). The aim of this paper is to describe development in these toddlers' ways of using and understanding counting words, as one representation of numbers, while participating in these interventions. Task-based interviews were conducted on five occasions; before, during and after the intervention. These provided an overview of the development of number knowledge among the toddlers. In this paper we specifically focus on the toddlers' learning of the the meaning of counting words. The research question we aim to answer is: How can toddlers' learning of counting words be described in terms of increase in discerned aspects of numbers?

Knowledge of numbers and counting words

A large body of research has given us a good understanding of children's general development of number knowledge within which ordinality, cardinality,

Camilla Björklund, University of Gothenburg
Hanna Palmér, Linnaeus University

part–whole relations, and representations have been shown to be essential aspects (Baroody & Purpura, 2017; Björklund et al., 2021; Fuson, 1992). There are however few studies involving toddlers as most tools for investigating numbers are based on children having knowledge of the spoken number system.

Counting words include several significant aspects necessary to discern for making use of them, and their meaning is also connected to the context in which they are used (Fuson, 1992). The meaning of ordinality implies that every object or counting word has an exclusive position in a sequence and relates to the others in the same sequence (Fuson, 1992). Cardinality then implies knowing that counting words may represent a set or composed unit of items (Baroody & Purpura, 2017). Cardinality is by Sarnecka and Carey (2008, p. 665) described as a ”principle stating that a numeral's cardinal meaning is determined by its ordinal position in the [counting] list”. This knowledge has been shown among most four-year-olds, and is a fundamental insight if one is to be able to make use of counting words and handle numbers in problem solving involving quantities. Cardinality is often observed in counting acts in which the last uttered counting word includes all the counted items (Gelman & Gallistel, 1978). Nuñes and Bryant (1996) do however raise some concerns regarding this interpretation of the cardinality principle, since it may reflect a learnt procedure rather than numerical understanding. Cardinality is furthermore a prerequisite for understanding numbers' part-whole relations, which allows children to compare sets, add, subtract and in different ways operate with numbers to solve numerical problems (Venkat et al., 2019).

Because of the abstract nature of numbers, they are only accessible through representations like spoken language, symbols and images (Duval, 2006). According to Lesh et al. (1987), learning is reflected in the ability to make connections *between* and *within* the representation used. A study by Gibson et al. (2019) has shown that children who use several representations, whether or not they are used in correct correspondence with a number of objects, more easily learn the cardinal meaning of numbers. In many situations, at least two representations are explicitly or implicitly used, depending on the mathematical activity (Duval, 2006), with pictures, verbal and written symbols, manipulatives, and real-world situations being representations that are often elaborated on in early mathematics education (see Lesh et al., 1987). Learning about numbers thus implies discerning their meaning (including ordinality, cardinality, and part–whole relations) mediated through different representations. This discernment does not appear on its own, but through communication with others (van Oers, 2010). Using several modes of representation in this communication is not enough either; the representations must be pointed out and demonstrated to the child, as neither representations themselves nor connections between and within them are automatically realized by children (Björklund & Palmér, 2022, 2023; Palmér & Björklund, 2023). Learning to use counting words is thereby a complex endeavor that goes beyond reciting the counting sequence, but should be possible to promote in early years because of the frequent use of counting words in daily communication and education.

Theoretical framework

The theoretical framework for this study is Variation theory of learning as this framework directs attention to how a learner experiences the meaning of a phenomenon (Marton, 2015). How a person experiences the meaning of numbers entails both a holistic experience (experiencing numbers "as" something), and a complex experience, in which discerned aspects of the phenomenon together constitute the meaning that the child attributes to it (Gibson & Gibson, 1955; Marton & Pong, 2005). This means, the child always has some way of experiencing numbers, based on previous and similar experiences. When the toddler discerns aspects of the phenomenon that he or she has not previously discerned, it changes the meaning for that toddler, leading him or her to a new way of understanding (experiencing) numbers. Through using variation theory it is possible to design an intervention offering children to experience necessary aspects of numbers through carefully selected patterns of variation and invariance (see Marton, 2015 and Björklund et al. 2021 for further details) and to make a detailed interpretation of what it means to learn something and by analysing learning outcomes in terms of changed ways of experiencing a phenomenon (see Björklund & Runesson Kempe, 2019). In every situation, several aspects of a phenomenon can be discerned, and those that are discerned are decisive for how the phenomenon is experienced. For example, if a child has not yet discerned the cardinal meaning of numbers, counting words are reduced to having the meaning similar to a nursery rhyme (the order of the words are stabile but bear no meaning of "numerical sets"). Learning then occurs when the learner discerns new and necessary aspects of that phenomenon, which in this study means a focus on those aspects that children have and have not yet discerned of counting words. Thus, the approach does not limit the analysis to "knowing or not knowing", but instead makes it possible to analyse *different ways of knowing*.

Method

The focus of this paper is on toddlers' knowledge of counting words. For three semesters, two researchers and three preschool teachers from three preschools in two Swedish municipalities collaborated. Activities focused on numbers were implemented in the preschools during these three semesters. The activities in the intervention were adapted to toddlers and differed in nature, for example, book reading, memory games, as well as dance and motoric play (se e.g., Björklund & Palmér, 2021; Palmér & Björklund, 2023). At the start of the study, the 27 children participating in the intervention were 12–27 months old. The children's guardians agreed to the participation of their children. The tools and methods of the study have been approved by the Swedish Ethical Review Authority (Dnr: 2019-01037). To investigate the advancement of toddlers' number knowledge, task-based interviews were conducted on five occasions: before, three times during and then after the intervention (Björklund & Palmér, 2021). In the first interview the mean age of the children were 1 year 5 months, in the second 1 year 9 months, in the third 2 years 1 month, in the forth 2 years 7 months and finally in the last interview, 3 years 1 month.

To overcome the challenge of toddlers' often non-verbal communication, the task-based interview was play-oriented, not depending on verbal responses (Björklund & Palmér, 2021). The interview consisted of seven tasks, framed in a narrative familiar to the children (i.e., a birthday party), with numerical aspects available to be discerned and reasoned about by the toddlers in different ways, for example by using words, gestures, and other actions. Each task included five levels of advancement, to offer the children opportunities to show their individual potential regardless of their biological age and to be used for longitudinal analyses of advancement. All interviews were conducted by the preschool teachers, who were trained in conducting research interviews with young children (Palmér & Björklund, 2022).

Analysis

The analysis is based on 60 hours of video-documented interviews, coded in NVivo. In this paper, the focus is on toddlers' use of counting words which was observed in tasks where the toddlers are enumerating or reasoning about "how many" questions. In accordance with variation theory (Marton, 2015), how the toddlers acted (made use of counting words) was considered to be a function of how he or she experienced the meaning of the counting words. Discerned and undiscerned aspects were central in the analysis because what is discerned constitutes the meaning the phenomenon has for the toddler in a particular situation.

First, the toddlers' ways of using counting words in the different tasks were categorized. These ways were, in line with variation theory, interpreted in terms of discerned aspects. This first part of the analysis ended up in a qualitative description (six categories) of the toddlers' knowledge of counting words. Second, we sorted the observed instances of the qualitative categories covering all the five interview occasions for all of the participating children. Through this procedure, four distinct learning trajectories emerged from the data set. In the results, examples from four toddlers illustrate these learning trajectories.

Results

The results are presented in two parts. First, the six qualitative different categories of ways the toddlers use counting words are presented, second four tables that illustrate the learning trajectories found among the toddlers in the study.

The qualitative different ways of using counting words are below described based on what aspects are discerned and not yet discerned.

A. Repeating words in a systematic chanting manner

Example: Pointing at each cookie while saying ”di, di, di”

When toddlers repeat words in a systematic chanting manner, they discern objects as differentiated units constituting a collection that can be demarcated with verbal and gestural actions. This is a pre-requisite for experiencing a set as partitioned and as a composite whole at the same time. However, the toddler has not yet discerned cardinality or ordinality of numbers. The way the toddler demarcates one word for each item in a collection indicates, however, some sense that words or certain utterances spoken in a rhythmic way may represent the items in the collection.

B. Random counting word

Example: Answering ”three” to every task regardless of the number of objects

The use of random counting words means that the toddlers discern counting words as a special kind of words that are to be used in certain situations, such as when asked ”how many”. Counting words are in this sense experienced as a representation for quantities. The use of random (or one and the same) counting word seems on the other hand to lack connection to the set of objects and thereby lack the meaning of cardinality or ordinality or any relations between numbers.

C. Reciting the counting rhyme unsystematically

Example: Answering ”one, two, five, eight” when asked to determine the number of three objects

Toddlers in the study sometimes recite the counting rhyme unsystematically, which is interpreted as they are discerning counting words as a collection of words with several differentiated examples (different words). The words do have some kind of representational meaning, because the words are used in a certain setting and in a certain way to answer the question ”how many”. However, the toddler does not yet discern ordinality because the string of words are random and the words do not seem to be related to one another other than as parts of a similar category of words. Neither have the words the meaning of cardinality because one single counting word is not used to determine the quantity of a collection of objects.

D. Reciting the counting rhyme in correct order, not connected to the set of objects

Example: Counting out loud, sometimes pointing with index finger at objects but does not stop counting when reaching the last object.

The toddlers may recite the counting rhyme in correct order but not connect the words to a set of objects, which means they discern ordinality in the sense that the counting words are ordered in a sequence, starting from ”one”. There is however no cardinality meaning discerned because the counting action does not end up in a collection answered by a single counting word. Likewise, toddlers who name each item (in the correct order) but are not flexble in their use (that ”number two” can be any object depending on where you start the counting act) have not discerned the cardinality of the counting words either. Similar as in the previous category, the words have a representational meaning, because the words are used in a certain setting and in a certain way as a ”string of words”.

E. Reciting the counting rhyme in correct order, stops at last counted object

Toddlers who recite the counting rhyme in the correct order often accompany this with pointing actions and conducts the counting in one-to-one correspondence. Ordinality can thereby be interpreted as discerned, especially if the toddler reacts to someone else's incorrect counting and/or pointing. If the child stops his or her counting action (words and gestures) on the last object, coordinating the reciting and pointing, this may indicate that cardinality is discerned. The counting words have a representational meaning in that they are coordinated with objects that are pointed at and uttered in a stable order, thereby representing a sequence and the ordinality of the counting words are foregrounded.

F. Correct counting word for a set of objects

> Example: ”Soon there will be five”, the toddler is building a tower of bricks one at a time, stops and points and counts when there are five bricks in the tower.

Some toddlers use (correct) counting words for sets of objects. Cardinality is then discerned as an aspect of number. Some are also able to make judgements of whether there are too many or too few objects to make a specific quantity that is asked for (with a counting word). Thus, among these toddlers, ordinality is also discerned. Representations of numbers expressed in the counting words also have a distinct meaning related to the cardinality and ordinality, to the toddlers.

Overview of discerned aspects

The above fine-grained analysis of the toddlers' way of using counting words provides a complementary view on number knowledge. This is shown for example when it comes to discerning the meaning of representations. Representations appear as a foundational aspect that foregrounds ordinal and cardinal aspects, which in turn are decisive for experiencing a numerical meaning of counting words (and thereby making possible to use counting words to enumerate and determine a quantity). The analysis further shows that representation in itself constitutes several aspects that determine its meaning, and use. Counting words seen as a representation of numbers can mean a certain kind of words to be used in certain kinds of situations. That is, a general idea of what the counting words are meant to be used for. When toddlers use counting words as a sequence (either random or in a stable order) they have differentiated the words from each other as examples of the same kind of representation. To understand representations in even more advanced ways, other aspects such as ordinality and cardinality are necessary to discern as well, in order for counting words to represent a numerical meaning.

Progress in development of counting word knowledge

Taking starting point in the above described categories of different ways of knowing and using counting words, we can conclude that there is a distinct difference in *how* but also *when* the toddlers make use of counting words. To further answer the research question, four types of learning trajectories were identified. The numbers in tables 1–4 indicate the number of observations in respective interview.

Linear progression in discerning necessary aspects

A typical progression is linear, which means that a toddler in the early interviews use counting words in ways that are interpreted as expressions of not discerning cardinality or ordinality. The toddler rather uses counting words and rhythmic chanting as certain kinds of actions and words to be used in certain situations. In the later interviews more necessary aspects are discerned, which is indicated by an increase of reciting of the counting sequence is in correct order and also connected to the set of objects to be counted (see table 1).

Table 1. *Example of linear progression in discerning necessary aspects*

Ways of using counting words	int 1	int 2	int 3	int 4	int 5
A. Repeating words in a systematic chanting manner	0	3	7	1	1
B. Random counting word	0	0	3	8	1
C. Reciting the counting rhyme unsystematically	0	0	0	5	0
D. Reciting the counting rhyme, correct order, not connected to the set of objects	0	0	0	2	7
E. Reciting the counting rhyme in correct order	0	0	0	2	15
F. Correct counting word for a set of objects	0	0	0	12	10

"Ketchup effect"

Some of the toddlers do not verbally express any discerned aspects of counting words in the first interviews. In the later interviews, however, they suddenly make use of counting words in ways that indicate their discerning all necessary aspects of counting words. Interestingly, for these toddlers all categories usually appear in parallel which may indicate some uncertainty about how to use counting words in different tasks (see table 2).

Table 2. *Example of "ketchup effect" in the discernment of necessary aspects*

Ways of using counting words	int 1	int 2	int 3	int 4	int 5
A. Repeating words in a systematic chanting manner	0	0	0	0	1
B. Random counting word	0	0	0	4	6
C. Reciting the counting rhyme unsystematically	0	0	0	1	2
D. Reciting the counting rhyme, correct order, not connected to the set of objects	0	0	0	4	5
E. Reciting the counting rhyme in correct order	0	0	0	5	8
F. Correct counting word for a set of objects	0	0	0	3	16

Even progress in discerning necessary aspects

A third observed learning trajectory shows an even progress in discerned aspects, where some actions in the early interviews indicate an awareness of ordinality and cardinality. However, the number of observations indicating this awareness increase, showing an even spread across the categories with tendencies to make use of counting words in a numerical manner more often in later interviews, which indicates their discerning both ordinality and cardinality in addition to representations (see table 3).

Table 3. *Example of even progress in discerning necessary aspects*

Ways of using counting words	int 1	int 2	int 3	int 4	int 5
A. Repeating words in a systematic chanting manner	0	0	0	0	0
B. Random counting word	0	4	3	2	4
C. Reciting the counting rhyme unsystematically	0	2	1	1	0
D. Reciting the counting rhyme, correct order, not connected to the set of objects	1	1	3	1	3
E. Reciting the counting rhyme in correct order	2	0	7	12	12
F. Correct counting word for a set of objects	0	3	9	8	15

No or limited discernment of necessary aspects

In a few cases, we observed a very limited use of counting words across all interviews indicating that the toddlers were not (yet) discerning any of the necessary aspects of counting words to make use of them in the numerical tasks (see table 4). Because no later interviews were conducted, it is not possible to draw any conclusions whether they are extreme examples of "ketchup effect" or if there might be other reasons for the toddlers not making use of counting words.

The observed progress in development of counting word knowledge among the toddlers in the study shows four typical learning trajectories. These are

Table 4. *Example of no or limited discernment of necessary aspects*

Ways of using counting words	int 1	int 2	int 3	int 4	int 5
A. Repeating words in a systematic chanting manner	0	0	0	0	0
B. Random counting word	0	0	0	0	0
C. Reciting the counting rhyme unsystematically	0	0	0	1	0
D. Reciting the counting rhyme, correct order, not connected to the set of objects	0	0	0	0	0
E. Reciting the counting rhyme in correct order	0	0	0	0	0
F. Correct counting word for a set of objects	0	0	0	0	0

different regarding at what time some of the necessary aspects of counting words are discerned for the first time, however, most of the toddlers (regardless of learning trajectory) show their discerning all necessary aspects in the fourth interview (as approximately 2.5 year-olds) and even more so in the last interview (with the exception of the fourth trajectory where no or very limited use of counting words have been observed). What the first three learning trajectories seem to have in common, is that all ways of using counting words, even those that do not indicate discerning ordinality or cardinality, are observed also in the later interviews. This might be due to the tasks used in the interviews, which the current analysis has not taken into account.

Discussion

The results show a frequent use of counting words among the toddlers but also a distinct difference in how toddlers make use of counting words. Studying this in terms of discerned aspects of counting words is an approach that complements most earlier research on young children's number knowledge, where focus often is set on the skills to solve a numerical problem that the children express when encountering numerical tasks (for example if they are able to produce a specific number of items when asked ”can you give me n”, see e.g., Sarnecka & Carey, 2008).

First, one result is that to learn the meaning of counting words is a very complex endeavor why what it means to ”know” numbers should be given more attention. For example, our study unfolds that to discern cardinality and ordinality it is necessary to discern the meaning of number representations, while discernment of cardinality and ordinality liberates a more advanced understanding of representations.. This might add an explanation to Gibson et al.'s (2019) study showing that children who use several representations, correct or incorrect in numerical sense, more easily learn the cardinal meaning of numbers. Thus, representations are themselves a complex learning object, closely related to other aspects of numbers. More studies, both theoretical and empirical, are needed to distinguish what constitutes this relationship between and within aspects of counting words.

Second, when taking a longitudinal perspective on the discernment of necessary aspects, we conclude that there are quite different learning trajectories among the toddlers. This is an important finding, because the toddlers have participated in the same interventions during a considerably long time period (three semesters). This finding has significant impact on our understanding of how to assess the outcome of interventions implemented in authentic preschool practices, but also on the theoretical understanding of children's numerical development. This study is however small scale and the participating children are not (yet) followed up in later ages; issues that should be considered and investigated further in future studies.

Third, at the same time as this study shows that it is possible to make fine-grained analyses of toddlers' actions as expressions of different ways of understanding counting words, the analysis reveals the necessity to include several tasks when drawing conclusions about discerned aspects. Otherwise observations of children counting in a one-to-one correspondence, item-to-counting word, stopping at the last counted item, may be falsely interpreted as the child understanding the cardinality meaning of counting words (see Gelman & Gallistel, 1978; Nuñes & Bryant, 1996).

Acknowledgements

The study was financed by the Swedish Institute for Educational Research under Grant no. 2018-00014.

References

Baroody, A. & Purpura, D. (2017). Early number and operations: whole numbers. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 308–354). NCTM.

Björklund, C., Ekdahl, A.-L. & Runesson Kempe, U. (2021). Implementing a structural approach in preschool number activities. Principles of an intervention program reflected in learning. *Mathematical Thinking and Learning*, 23 (1), 72–94. doi: 10.1080/10986065.2020.1756027

Björklund, C., Marton, F. & Kullberg, A. (2021). What is to be learnt? Critical aspects of elementary arithmetic skills. *Educational Studies in Mathematics*, 107 (2), 261–284. doi: 10.1007/s10649-021-10045-0

Björklund, C. & Palmér, H. (2021). Designing a tool for exploring toddlers' number knowledge in preschool. In Y. Liljekvist, L. Björklund Boistrup, J. Häggström, L. Mattsson, O. Olande & H. Palmér (Eds.), *Sustainable mathematics education in a digitalized world. Proceedings of MADIF 12* (pp. 1–10). SMDF.

Björklund, C. & Palmér, H. (2022). Teaching toddlers the meaning of numbers: connecting modes of mathematical representations in book reading. *Educational Studies in Mathematics*, 110, 525–544. doi: 10.1007/s10649-022-10147-3

Björklund, C. & Palmér, H. (2023). Enhancing Swedish toddlers' learning opportunities through interactions with pictures and narrative designed for numerical learning purposes. *Early Childhood Education Journal* (Advance online publication). doi: 10.1007/s10643-023-01556-x

Björklund, C. & Runesson Kempe, U. (2019). Framework for analysing children's ways of experiencing numbers. In U. T. Jankvist, M. Van den Heuvel-Panhuizen & M. Veldhuis, (Eds.), *Proceedings of CERME11*. Freudenthal Group & Freudenthal Institute, Utrecht University and ERME.

Duval, R. (2006). A cognitive analysis of problems of comprehension in a learning of mathematics. *Educational Studies in Mathematics*, 61 (1), 103–131.

Fuson, K. (1992). Research on whole number addition and subtraction. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 243–275). Macmillan.

Gelman, R. & Gallistel, C. (1978). *The child's understanding of number*. Harvard University Press.

Gibson, D., Gunderson, E., Spaepen, E., Levine, S. & Goldin-Meadow, S. (2019). Number gestures predict learning of number words. *Developmental Science*, 22:e12791, 1–14. doi: 10.1111/desc.12791

Lesh, R., Post, T. R. & Behr, M. (1987). Representations and translations among representations in mathematics learning and problem solving. In C. Janvier (Ed.), *Problems of representation in the teaching and learning of mathematics* (pp. 33–40). Lawrence Erlbaum.

Marton, F. (2015). *Necessary conditions of learning*. Routledge.

Nuñes, T. & Bryant, P. (1996). *Children doing mathematics*. Blackwell Publishers.

Palmér, H. & Björklund, C. (2022). Divergence in interviews with children: improving research quality. *Designs for learning*, 14 (1), 52–57. doi: 10.16993/dfl.187

Palmér, H. & Björklund, C. (2023). The teaching of numbers in common preschool activities: a delicate balancing act. *Early Childhood Education Journal*, 51, 971–980. doi: 10.1007/s10643-022-01354-x

Sarnecka, B. & Carey, S. (2008). How counting represents number: what children must learn and when they learn it. *Cognition*, 108 (3), 662–674.

Van Oers, B. (2010). Emergent mathematical thinking in the context of play. *Educational Studies in Mathematics*, 74, 23–37.

Venkat, H., Askew, M., Watson, A. & Mason, J. (2019). Architecture of mathematical structure. *For the Learning of Mathematics*, 39 (1), 13–17.

When implementing the method Thinking classroom, the didactical contract is hard to break

Maria Lundkvist and Kerstin Larsson

Efforts have been made to change mathematics teaching from direct instruction towards methods where the students are enticed to think mathematically. This has proven difficult to achieve, partly because of classroom norms and conflicting didactical contracts. *Thinking classroom* is a teaching method that has been suggested to tackle these challenges. This study investigates the relationship between didactical contract and Thinking classroom. For this purpose, one upper secondary school teacher was invited to implement the method in their teaching. The teacher's introductions of tasks to the students were analysed in relation to Thinking classroom's recommendations and the pre-existing didactical contract. Our study shows that although Thinking classroom is designed to bypass norms the didactical contract is hard to break.

For a long time and in many countries, efforts have been made to change mathematics teaching from direct instruction towards, what we call "students' active learning" (Schoenfeld & Kilpatrick, 2013). These ideas have for instance been labelled as problem-based learning, inquiry-based learning, and reform teaching. Even though these efforts have been underpinned by slightly diverse theoretical ideas (Merritt et al., 2017), there are also many similarities. Problem-solving and mathematical reasoning have been central and one of the purposes has been to foster students to think and reason by themselves actively. However, in changing one's teaching, teachers may be hindered by different classroom norms (Hofmann & Ruthven, 2018).

In addition to classroom norms, a mathematical classroom also includes a *didactical contract* (DC) which constitutes what is expected of both the teacher and the students (Brousseau, 1997). In direct instruction, the students expect the teacher to show how the tasks are solved and the teacher expects the students to listen and later use the solving method. A DC is implicit and becomes visible only when the teacher or the students contradict the expected. The DC in direct instruction can also be deliberately broken, of which *Thinking classroom* is an example (Liljedahl, 2022).

Maria Lundkvist, Luleå University of Technology
Kerstin Larsson, Luleå University of Technology

The main reason for using Thinking classroom (TC) is to compel the students to think mathematically by themselves (Liljedahl, 2022). This teaching method consists of several practices with the salient features that the students work in groups that are formed randomly and standing by vertical non-permanent surfaces (for example a whiteboard). The teacher avoids direct instruction by, for example, not answering students' questions on how to solve the task. Therefore, TC aligns with efforts to change mathematics teaching toward students' active learning. Thus, we suggest that this prescriptive method, described in Liljedahl (2022), might be considered as a tool for teachers who wish to develop their teaching in this direction.

Background

First, we describe elements of Thinking classroom relevant to this study as a theoretical background for this paper. Thereafter, we briefly present the theoretical concept of didactical contract.

Thinking classroom (TC) is a method for teaching mathematics that has been studied and developed through a series of smaller studies (Liljedahl, 2018, 2019; Liljedahl & Allan, 2013). In short, TC consists of 14 practices that together can be described as a way to compel students to think mathematically by themselves rather than to mimic the teacher or the teaching material when learning mathematics (Liljedahl & Allan, 2013). The first three practices concern (1) how to choose tasks, (2) use randomised grouping, and (3) that the students are standing and working on erasable surfaces.

The first practice deals with choosing tasks that should compel the students to think mathematically. In the first lessons following TC, the tasks should be non-curriculum [1] (Liljedahl, 2022). This means they are supposed to be a new type of task that is unfamiliar to the students. Furthermore, the tasks should be engaging and easy enough so that every student can begin solving them, while also allowing more complex reasoning. Moreover, to keep the students focused the task should not take more than five minutes for the teacher to introduce. An example of such a non-curricular task is to write expressions using four fours and arithmetical operations that equal for example all numbers 1–20 (Boaler, 2011). This task is easy to explain to the students and for them to begin with but requires creative thinking so that all the expressions can be found. Later, when the teacher and the students have gotten used to TC, the tasks are changed to curricular tasks that need to be covered in the course (Liljedahl, 2022).

The second practice, randomised grouping, implies dividing the students into groups who work together during the lesson (Liljedahl, 2022). The randomisation is to be done openly, that is, the students can see that they really are assigned to a group randomly and not by the teacher. The randomised grouping is done separately for each lesson to avoid students taking specific roles in a fixed group, which might prevent them from thinking freely. The recommended size is three students in each group and if not possible then two. More than three students might lead to smaller groups within the group. The risk of having only two students is that the knowledge variation is not wide enough, or that it is too wide, to engage in fruitful discussions.

The third practice is to have the students work on vertical non-permanent surfaces for example, a whiteboard. The vertical surfaces force the students to stand up, which raises energy levels and supports cooperation since acting passively when standing is more difficult (Liljedahl, 2022). A non-permanent working area has been shown to encourage students to start working on the task faster and more freely since they know that everything easily can be erased. Moreover, to promote cooperation between the students, each group should only have one marker (Liljedahl, 2018).

The first three practices are recommended to be implemented together and the remaining eleven practices[2] can be implemented one by one. The reason for this is that neither of the first three practices is enough to break classroom norms by themselves (Liljedahl, 2022). These three practices function as a signal that the mathematical work is going to be something they are not used to, which implies that the students more freely can step outside their normal roles in a mathematics classroom. We view Liljedahl´s practices, more specifically the first three, as a theoretical background for this paper.

As a theoretical framework, we employ the concept of didactical contract (Brosseau, 1997), which can explain some of the struggles a teacher can face when trying to change the teaching method. A breach of the didactical contract (DC) has been used in other studies, for instance as an explanation for students' difficulties with mathematical modelling (Jankvist & Niss, 2020). A DC is formed in every mathematical classroom even though it is not explicitly expressed (Brousseau, 1997). Since the DC is implicit, it becomes visible only when one part breaks the contract in some way. Moreover, it is a social agreement that both the teacher and the students are reciprocally responsible for upholding. For example, when the teachers present a task or model on how to solve a specific type of task, in Swedish called ”genomgång”[3], it is often made by posing questions to the students (Andrews & Larson, 2017). These questions are often rhetorical and/or closed and seem to be a part of the DC (e.g. Jankvist & Niss, 2020).

Since the DC is mostly implicit the teacher might not be aware of its existence and therefore becomes a hurdle for teachers when changing their teaching methods. If changing one's teaching involves new demands, such as coaxing students to think rather than mimic, students might react negatively. Not necessarily because they do not want to think by themselves, but because they react to the change in itself. In addition, a teaching method such as TC, that engages students to think actively, is more time-consuming (Serrano Corkin et al., 2019). This might conflict with demands from the curriculum to engage in every mathematical area, thus forming further obstacles to changing one's teaching.

This paper examines what a teacher can experience when introducing a task while using the first three practices of Thinking classroom. We follow one mathematics teacher when using this teaching method in a Swedish upper secondary classroom. The following questions guide our study: 1) How does the teacher introduce tasks to her students while following the method of Thinking classroom? 2) Which part of the pre-existing didactical contract can interfere with the intentions of Thinking classroom during the introduction?

Method

First, we present the participants, including the researcher's cooperation with the teacher. This is followed by a description of the pre-existing DC by illustrating how ordinary lessons usually start. Thereafter, data collection and analyses are described and explained. The data presented here comes from a pilot study, conducted by the first author, as a part of a PhD project that implements TC at upper secondary schools.

The researcher presented the project to a group of teachers at one upper secondary school (ages 15–19) and one teacher volunteered to participate. The teacher was experienced with more than ten years working at this school. All 16 students in the group chosen by the teacher gave their consent to participate in the study. The cooperation with the teacher involved two individual meetings where the researcher presented the main idea behind TC and which tasks to use in the six consecutive lessons were discussed. The first three lessons built upon non-curricular tasks chosen from Liljedahl's book (2022) and the following on growing patterns, both 2D and 3D. Each lesson revolved around one task, thus six tasks in total were chosen.

The pre-existing DC in this classroom during the beginning of ordinary lessons comprises students' individual work with a task from the textbook, followed by pair work on the same task, which the teacher then solves in front of the students. Thus, the students might or might not work on the task, knowing that they will soon get the solution, which then can be mimicked. The solution in front of the class is often followed by a genomgång where the teacher presents something new, which both can entice students to mimic. The pairs are set by the teacher and only change if the cooperation does not work sufficiently. This implies that the pre-existing DC includes that the teacher can decide who the students are to cooperate with and that mimicking solutions is common. The teacher frequently poses questions to the students during both the presentation and when showing the solution, indicating that questions from the teacher are part of the pre-existing DC. This scenario was described by the teacher and some of the students in interviews and confirmed by fieldnotes.

Since we were interested in visible features such as the teacher's and students' gestures and body language, as well as verbal utterances, video recording was conducted, enabling a multimodal analysis (Jewitt, 2012). The six lessons when the teacher used TC, were all filmed as a whole except for lesson five where the recording started after the teacher's introduction. For this paper, the introduction parts of the video recordings of lessons 1–4 and 6 were analysed.

To understand and explain the findings from the analyses the theoretical construct of DC was used. The reasons for this choice were a) the explicit recommendation of Liljedahl (2022) that it is important to create a breach of classroom norms when implementing TC and b) that the DC is unspoken and reciprocally upheld by both teacher and students thus, difficult to change (Jankvist & Niss, 2020). Both visible breaks of the pre-existing DC, and when it overruled practices of TC became apparent by this theoretical framework. These choices were underpinned by our aim to examine what a teacher can experience when implementing the first three practices of TC.

We analysed the length of the introduction, the students' positions, the arrangements of the classroom, and the randomisation process since these aspects occur during the introduction. We wanted to investigate what, if any, of TC practices created breaches of the DC visible through these aspects of the introduction, since one rationale behind TC, according to Liljedahl (2022), is to break classroom norms. The data also contained other aspects that were not obtained by this analysis but were salient to us. We found these aspects to be either breaking or confirming the pre-existing DC, hence assisting or challenging the implementation of TC. The following categories, created by inductive content analysis (Krippendorff, 2019) captured these salient aspects: the teacher's use of a manuscript, questions asked by the teacher, and questions asked by the students. Questions asked by teachers in a mathematics classroom can be categorised in different ways. A common pattern of classroom questions is IRF – initiation, response and feedback – which might impede mathematical thinking and reasoning (Attard et al., 2018). In this paper though it is more valid to note how often the teacher asks questions than to categorise them since most of them were closed and rhetorical and were part of the pre-existing DC.

Results

Firstly, an overview of all five videotaped introductions is given. After that, we paint a picture of the teaching situation by providing a thick description of the introduction of lessons 1 and 4 which are chosen as examples of using both non-curricular and curricular tasks. Throughout the results, we demonstrate our conclusions from the analyses by connecting the result with recommendations from TC and/or pre-existing DC.

Overview of the introduction

An overview of introductions in lesson 1–4 and 6 is shown in table 1. The time used for the introduction of the tasks ranged from 1:10–6:06 where the two longest were the first two introductions. In TC the recommended timeframe for

Table 1. *Overview of the introductions*

Introduction number	1	2	3	4	6
Length of introduction min:sec	05:32	06:06	03:18	03:01	01:10
Number of questions from the teacher	23	15	20	4	3
Number of questions from students	0	4	0	1	0
Looking at manuscript (% of time)	11,7 %	8,7 %	12,1 %	33,7 %	-
Number of times looking at the manuscript + asking questions to the researcher	14 + 2	15 + 2	10 + 0	9 + 0	- + 0

the introduction is a maximum of five minutes which means that the first two exceeded that timeframe. During the introductions the teacher asked several questions, the greatest number of questions asked by the teacher was 23 in Introduction 1, after that the number of questions decreased. One interpretation of these results is that the prolonged introductions in the first two lessons were due to the use of more questions than in later. In Introduction 3 the teacher asked 20 questions and kept the timeframe which is explained by the fact that almost all questions were rhetorical, which the teacher answered herself. Another interpretation of the prolonged introductions is that the tasks are unfamiliar to the teacher. The few questions the students asked mainly concerned clarifications, for example, ”What does factor mean?”. Consequently, the students’ questions did not affect the length of the introductions.

The teacher had the task written on a paper in all lessons but one, hereafter referred to as a manuscript. In the first three introductions, the teacher looked at the manuscript around ten per cent of the time and one third in Lesson 4. Furthermore, she asked the researcher four questions in total during the first two introductions. These actions may be perceived as the teacher’s unfamiliarity with both tasks and teaching method.

Another feature concerning the introductions is the randomisation process, not presented in table 1. A deck of cards was used to randomly divide the students into groups of two or three students each. The researcher had the responsibility to count the number of students present and make sure that the number of cards matched to facilitate for the teacher. During all except one of the lessons, the teacher hands out the cards. No reactions from the students regarding the grouping were visible in the recordings. One possible interpretation is that the pre-existing DC is not breached by the randomisation.

Lesson 1

The researcher and the teacher prepared the classroom by organising furniture and hanging the vertical non-permanent surfaces (VNPS), brought by the researcher, in strategic places, which took about ten minutes. When opening the door 15 students entered the classroom. They looked at all the desks and chairs that were cramped together to make space around the VNPS (see figure 1)

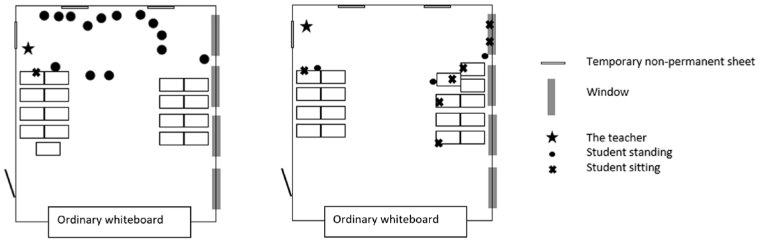

Figure 1. *A sketch of where the teacher and the students were located in the classroom in Lesson 1, to the left, and Lesson 4, to the right*

and left their bags on top of the tables. The teacher stood by one of the VNPS at the back, holding some paper and waiting for the students to gather around her, which they did. Almost all students stood up (see figure 1) and looked towards the VNPS where the teacher presented the task. The fact that almost all students stood up and how the classroom was organised is a breach of their pre-existing DC. During the next five and a half minutes the teacher explained the task. Throughout that period the teacher asked 23 questions (see table 1) mostly to engage the students, e.g.:

01:10	Teacher:	What number do you want to use?
01:14	Student 1:	Three
01:19	Teacher:	And what operation do you want to use?
01:23	Student 2:	Well ... times
01:30	Teacher:	What other number do you want to use?
01:32	Student 1:	Five
01:36	Teacher:	And ... what is that?
01:37	Student 1:	Fifteen

The use of questions during the introduction can be interpreted as the teacher's ingrained behaviour and a part of the pre-existing DC. This prolongs the introduction and thereby contradicts the recommendations of TC.

The questions asked by the teacher to the researcher concerned the task itself and when to form the groups. The teacher looked at her manuscript 14 times, equivalent to just over ten per cent of the introduction. As mentioned before this shows that TC is something unfamiliar to the teacher and is illustrating a breach of the pre-existing DC. After the introduction, the researcher showed the students that the workspaces were numbered and handed the cards to the teacher who gave them to the students. The researcher thereby intervenes and thus may override the pre-existing DC. Upon being given a card each student walked to their workspace and started working on the task. There can be several reasons for the students' lack of reaction, connected to TC, the pre-existing DC, and the researcher's presence, further discussed later.

Lesson 4

Once again, the teacher and the researcher arranged the classroom before the students entered. The teacher started the introduction by drawing a 2D pattern on one of the temporary VNPS while she described what she was doing to the students. After one and a half minutes the drawing of the first two figures of the pattern was finished whereas the teacher explained the purpose of the task. The manuscript was looked at nine times for approximately one third of the introduction, most time spent after the drawing was done. Four questions were asked by the teacher, all of them rhetorical. Most of the students sat on benches or a windowsill (see figure 1) while looking towards the teacher. At the end, one student asked the teacher to repeat the purpose of the task, no other questions were asked by the students. When the introduction was finished approximately three minutes later, the researcher handed out the cards to the students.

There are three differences between this introduction compared to the one in Lesson 1. Firstly, the recommended timeframe is held which might indicate that the teacher has adapted to TC, or the DC has changed. Secondly, the teacher used the manuscript in a larger proportion of the introduction. This is hard to explain through TC or DC, but one assumption is that the teacher wanted to ensure that the purpose of the task was accurate. Thirdly, most students sat down which placed them further away from the teacher. Our interpretation is that the pre-existing DC overruled the recommendations of TC.

Discussion

Even though many components of Thinking classroom (TC), at a surface level, seem easy to implement as described in Liljedahl (2022), they might involve a more profound change of pre-existing didactical contract (DC). This is suggested to be a cumbersome process (Jankvist & Niss, 2020). A breach of the DC is not only challenging for the students but also for the teacher. As shown in the results some aspects of the pre-existing DC might be ingrained habits that contradict the effort to breach the DC.

For instance, the time spent on the introduction was prolonged by the teacher's habit of asking questions. According to Andrews and Larson (2017), questions posed by the teacher during a genomgång are very common and applied to this teacher. During the interview, the teacher emphasised that she wants the students to answer her questions to keep the students engaged. However, as shown in the excerpt it can be time-consuming even though the predominant questions were closed and followed the IRF-pattern (Attard et al., 2018), without verbal feedback. We suggest that the pre-existing DC concerning genomgång and questions collide with Liljedahl's (2022) recommendation regarding the length of the introduction.

Over the sequence of lessons, more and more students chose to sit down during the introduction. During the introduction, Liljedahl (2022) recommends, in practice 6, that the students should stand near the teacher. In Introduction 1 all students but one did just that without any prompts from the teacher or researcher. We presume that the obvious breach of the DC when entering the classroom in Lesson 1, with all the furniture cramped together and the teacher standing at the back of the classroom, caused this. The teacher never encouraged the students to stand up as this recommendation never was presented to her. This demonstrates that it takes more than the first three practices of TC to uphold a new DC. In our opinion, it is important to consider how to present the task including the timeframe and the students' positions. Therefore, these recommendations ought to be incorporated into practice 1.

We suggest that the use of a manuscript indicates the teacher's unfamiliarity with the task and the teaching method. It is neither a recommendation in TC nor a part of the pre-existing DC. The amount of time spent looking at the manuscript prolongs the introduction which might influence the introduction to extend the recommended time of five minutes (Liljedahl, 2022). We find it surprising that an experienced teacher uses notes to this extent. However, this can be explained by the extraordinary situation. The tasks were new to her, the method was not only unfamiliar but also introduced briefly. On top of that, a researcher was present in the classroom to videotape the teacher.

One of the practices suggested by Liljedahl (2022) aiming for a breach of DC is randomised grouping. Therefore, one might expect students to react in some way. However, these students did not show any reaction visible in the recordings. We find three possible explanations for this. One is the fact that the researcher was present in the classroom and involved in the randomisation process, thus providing an alteration of the ordinary teaching. However, there are no other situations that can be linked to the presence of the researcher therefore we find this explanation unlikely. Another explanation might be that the first three practices in TC are designed to create a breach of the pre-existing DC. The classroom looks different, the teacher is standing in a different place, there are VNPS on the walls, they must stand, and there is a new kind of task. All this together might create such a breach of the pre-existing DC that the students just accept the randomisation. Although, if this was the case any objections would have been noticed in later lessons when the novelty of TC has worn off, as we can see in the fact that more and more students sat down instead of stood up. However, we connect the lack of reactions to the fact that this teacher uses peer work regularly where the students cannot choose whom to work with, hence a part of their regular DC.

Other challenges when implementing TC of a more practical nature are what kind of VNPS to use, where to hang them, and how to make room around them. In upper secondary school teachers and students normally change classrooms between lessons which entails that the time to rearrange the classroom is limited. We regard these practicalities to be important for schools to consider before implementing TC since it is challenging enough for teachers to change their teaching.

The chosen tasks have neither been presented in detail nor discussed. The reason for this is that the task itself is chosen in beforehand and therefore not a part of the introduction per se. In coming studies, it might be interesting to investigate what impact different tasks have on, for example, student engagement or reasoning possibilities.

We acknowledge that this study has limitations in its design which makes it hard to infer conclusions regarding all aspects of how a pre-existing DC can affect the implementation of TC. To analyse this further, we would need data from a series of whole lessons, including student work on the VNPSs, and more theoretical underpinning. We plan to further explore this phenomenon in coming studies. However, as a first glimpse of what challenges, and possible explanations of the challenges' nature, we find the results of interest to the research community. Moreover, our findings might support teachers and schools in implementing TC being prepared for the challenges and pursuing the changing of teaching methods.

In upcoming studies, we will provide teachers with a more in-depth introduction to TC in advance and preferably discuss the different practices with other teachers. Other takeaways concerned what kind of VNPS to use, how long it takes to arrange the classroom, and remembering to start the recording on time.

The main conclusion that we draw from this study is that even though the first three practices of TC seem easy to implement they are not; they involve much more than what is seen on the surface. For instance, the presentation of the task demands a more thorough preparation than can be perceived when reading the book by Liljedahl (2022).

References

Andrews, P. & Larson, N. (2017). Analysing genomgång: a Swedish mathematics teaching lesson event. *Nordic Studies in Mathematics Education*, 22 (3), 85–105.

Attard, C., Edwards-Groves, C. & Grootenboer, P. (2018). Dialogic practices in the mathematics classroom. In J. Hunter, P. Perger & L. Darragh (Eds.), *Making waves, opening spaces. Proceedings of the 41st MERGA conference* (pp. 122–129). MERGA.

Boaler, J. (2011). *Elefanten i klassrummet: att hjälpa elever till ett lustfyllt lärande i matematik*. Liber.

Brousseau, G. (1997). *Theory of didactical situations in mathematics: didactique des mathématiques, 1970–1990*. Kluwer Academic Publishers.

Hofmann, R. & Ruthven, K. (2018). Operational, interpersonal, discussional and ideational dimensions of classroom norms for dialogic practice in school mathematics. *British Educational Research Journal*, 44 (3), 496–514. doi: 10.1002/berj.3444

Jankvist, U. T. & Niss, M. (2020). Upper secondary school students' difficulties with mathematical modelling. *International Journal of Mathematical Education in Science and Technology*, 51 (4), 467–496. doi: 10.1080/0020739X.2019.1587530

Jewitt, C. (2012). *An introduction to using video for research*. National Center for Research Methods Working paper 03/12.

Krippendorff, K. (2019). *Content analysis*. SAGE.

Liljedahl, P. (2018). Building thinking classrooms. In A. Kajander, J. Holm & E. Chernoff (Eds.), *Teaching and learning secondary school mathematics: Canadian perspectives in an international context* (pp. 307–316). Springer. doi: 10.1007/978-3-319-92390-1_29

Liljedahl, P. (2019). Conditions for supporting problem solving: vertical non-permanent surfaces. In P. Liljedahl & M. Santos-Trigo (Eds.), *Mathematical problem solving: current themes, trends, and research* (pp. 289–310). Springer.

Liljedahl, P. (2022). *Att bygga tänkande klassrum i matematik*. Gleerups.

Liljedahl, P. & Allan, D. (2013). Studenting: the case of ”now you try one”. In A. M. Lindmeier & A. Heinze (Eds.), *Proceedings of PME 37* (pp. 257–264). Leibniz Inst. for Science and Mathematics Education.

Merritt, J., Lee, M. Y., Rillero, P. & Kinach, B. M. (2017). Problem-based learning in K–8 mathematics and science education: a literature review. *Interdisciplinary Journal of Problem-Based Learning*, 11 (2). doi: 10.7771/1541-5015.1674

Schoenfeld, A. H. & Kilpatrick, J. (2013). A US perspective on the implementation of inquiry-based learning in mathematics. *ZDM*, 45 (6), 901–909. doi: 10.1007/s11858-013-0531-5

Serrano Corkin, D., Coleman, S. L. & Ekmekci, A. (2019). Navigating the challenges of student-centered mathematics teaching in an urban context. *Urban Review*, 51 (3), 370–403. doi: 10.1007/s11256-018-0485-6

Notes

1 In Sweden all tasks can be labelled curricular since the students should learn problem-solving, mathematical reasoning, and communication. Although, we choose to follow the notion presented by Liljedahl (2022).

2 These eleven practices include for example how to furnish the classroom, how the teacher answers questions, and how the teacher evaluates.

3 No English word fully captures the meaning of the Swedish concept *genomgång*.

Preschool class students' discernment of number structure in a spatial pattern

Camilla Björklund, Jessica Elofsson,
Angelika Kullberg, Anna-Lena Ekdahl,
Ulla Runesson Kempe and Maria Alkhede

In this paper, we present what 361 preschool class students discern of number structures in a task containing a spatial pattern. The analysis, grounded in the variation theory of learning, focuses on the students' discernment of composite units as parts of a larger whole. To discern a task as a structure of parts and whole is presumed to be significant for what the student can do with numbers, e.g. to determine the exact number of objects in a spatial pattern. The results reveal four different ways in which students discern structures, each one with implications for the students' abilities to determine the number of the pattern. The results induce implications for the development of mathematics teaching in preschool classes and also raise further questions for research about students' arithmetic learning.

In the literature on early mathematics learning, there is great interest in understanding what skills children need to develop to make use of numbers for successful problem-solving. One such skill that a coherent research field (e.g. Baroody & Purpura, 2017) points out is understanding numbers as composite units, presumably leading to the ability to use addition and subtraction strategies based on number structures. Students develop their ability to see numbers as units by distinguishing relationships within and between numbers, for example, that seven consists of three and four and that seven is three more than four (see Baroody, 1987). Seeing numbers structured in such part-whole relations further means that a student can distinguish groups of objects in a set, connections between these and how they together form a whole that can be called a number (Venkat et al., 2019). The structure is thus the underlying regularity and general principle that defines numbers, which can be recognized, for example, in spatial patterns (e.g. a three-by-four array is seen as ”three groups of four” and also ”12”).

Camilla Björklund, University of Gothenburg
Jessica Elofsson, Linköping university
Angelika Kullberg, University of Gothenburg
Anna-Lena Ekdahl, Jönköping University
Ulla Runesson Kempe, Jönköping University
Maria Alkhede, University of Gothenburg

Students who do not see numbers as composed units and how such units can be combined to make larger units (in a part-whole relation) have shown to suffer difficulties because they cannot make use of powerful arithmetic strategies, such as making use of ”ten” as a benchmark and in particular handling multi-digit tasks (Ellemor-Collins & Wright, 2009; Neuman, 1987; Runesson Kempe et al., 2022). It is thereby necessary to gain more knowledge about young students' understanding of numbers and relations within and between numbers. To contribute to the field of research and knowledge about young students learning about numbers, we pose the research question: In what ways do preschool class students discern number structures to determine the number of a spatial pattern? By focusing on what students discern when encountering a specific task, in terms of number structures, we might get insights into what it means to discern structures that enables determining the number of the pattern in powerful ways. This can reveal important insights to what would be necessary to focus on and facilitate the learning of in early arithmetic education, to help students learn to discern number structures.

How to determine the number of a set

Many observations are presented of how children proceed in arithmetic skills development and which strategies they use in solving problems about quantities (Fuson, 1992). Sprenger and Benz (2020) describe the development path as children first not being able to distinguish any number structures that would help them determine the quantity of a set of objects. They then learn to distinguish parts or groups of objects but cannot yet connect these to one larger unit to determine the number of the whole set. Eventually, children are able to use number structures to determine the number of the set and usually solve these kinds of tasks in a proficient way. What remains unclear in the research, however, is what makes children able or unable to use number structures.

For several decades, studies have been conducted regarding cognitive processes of determining exact numbers, an intuitive process usually called (perceptual) subitizing (Kaufmann et al., 1949; Wynn, 1998). Regardless of how objects are arranged, most people have this innate ability to recognize small numbers up to three or four without counting. When numbers are larger, people can still determine exact numbers through a process called conceptual subitizing, which means to identify smaller sets as parts of a larger set, also without counting (Clements et al., 2019). The latter process depends on learned structures (e.g. seeing a group of three and a group of four, instantly knowing they constitute a set of seven together). The ability to see a set of objects as a composite unit has shown to be of great importance for students' continued mathematics learning (Hunting, 2003; Paliwal & Baroody, 2020), not least when the number range increases and tasks contain tens and hundreds or more (Ellemor-Collins & Wright, 2009).

Research and theories about children's development of counting skills are extensive but agree that a prerequisite is children's ability to discern units, that is, that objects in the outside world can be seen as belonging to a common collection or group, which is not initially numerical in its meaning but rather has the meaning that something can belong to a certain category of objects and constitutes an indefinite ”manyness” (Steffe, 1991). Experiencing patterns as numerical, on the other hand, means that it is the number of objects arranged in a certain pattern that appears to the child. This is a basis for distinguishing how groups can be seen as parts of a larger whole, that is, creating a structure.

According to Benz (2013), children can already in preschool age experience structures in quantities and use them to determine the number of sets. When students (perceptually) structure numbers (a whole) into subgroups, it implies a special ability to see wholes and parts. Schöner and Benz (2018) observed, for example, that when a student says, ”Two, three, and two more is seven”, the student discerns a structure based on number relationships in the form of parts that add up to 7 and can use this structure to determine the sum. It could be interpreted as the student discerning a group of two and a group of three that makes five, and five and two are seven, that is, several composite units that relate to a whole.

Previous studies have also focused on what kind of groupings students perceive. Sprenger and Benz (2020) investigated, with the help of technology that follows eye movements (eye-tracking), how 5-year-olds visually group objects when asked to determine numbers in a set. They found that the kind of grouping, and the use of structure, varied depending on the number of objects in the sets, that is, whether there were five, seven or nine objects arranged in a two-by-five array (egg carton), but also on how the objects were arranged in subitizable groups (three in one row and two in a row below, or as one row of five). Five objects placed in a row (in the egg carton) made the children use structure to a lesser extent than five objects arranged as groups of three and two. Mandler and Shebo (1982) also showed empirically that the way objects are arranged, and thus how they are perceived, matters for the certainty of determining numbers in a set.

Theoretical framework

The study we have conducted takes its starting point in Variation theory of learning (Marton, 2015). Variation theory conjectures that any action, such as how a student is determining a number in a spatial pattern, is a function of what the student discerns when encountering the specific phenomenon. When we approach for instance a mathematical task, some aspects are noticed, they come to the fore of our attention, and are related. How a student handles a task (reasoning about and indicating with words and gestures what he or she discerns) can thereby be interpreted as a result of how the awareness is structured at a particular moment and what is discerned in the task. Different responses observed among students who encounter the same task, including those having difficulties, are explained in terms of them not (yet) having discerned certain aspects of the task, in our case whether units involved and relations between and within them are attended to. Simultaneous discernment of several aspects is considered to be key in the progress in students' awareness because it liberates new ways of handling the task (this has been shown both theoretically in e.g. Marton, 2015, and empirically in e.g. Runesson, 2006; Björklund et al., 2021). This makes the theory useful for studying and interpreting qualitative different ways of encountering a task.

The study

The basis for the analysis is task-based interviews conducted with 361 preschool class students. The students are, on average, 6 years old the year they attend preschool class. The students were recruited from five different municipalities in the Southern parts of Sweden. Students' legal guardians had given their informed consent. In this particular study, our focus is on what structures are discerned in the spatial pattern, as they appear to the students in the total data set.

The interview consisted of tasks that the students were to answer orally. A protocol was used where the researcher made notes of the students' verbal answers and actions such as pointing, circling or gestures with fingers. Each interview took about 15 minutes. For the purpose of this paper, we have chosen one task for analysis: a spatial pattern where 12 objects are placed in 3 rows and 4 columns centred on an A4-sized paper [1]. The student is shown the pattern (see figure 1) and given two questions orally: a) "How can you see, how many dogs there are?" and b) "How many dogs are there, all together?". In the task, the number of objects is more than what cognitive processes such as subitizing include, which means that the student must use some kind of strategy to determine the number of the objects, such as distinguishing structures that appear

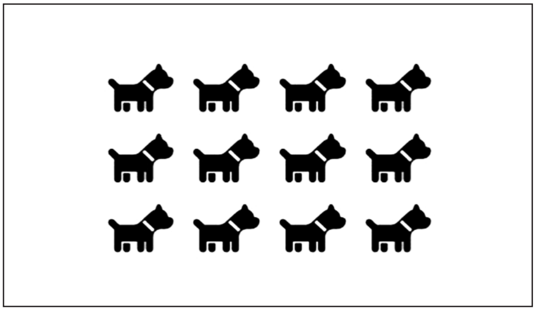

Figure 1. *Task with spatial pattern (11 x 7 cm on A4 size paper) to be answered orally*

in the pattern. The vast majority of the students (95 %) was able to complete the task (correct or incorrect) in one or another way, only 5 % did not complete the task or their answer was not possible to interpret.

Analysis

The students' responses to the tasks in the interview were documented in a protocol with a particular focus on their choice of strategies to solve the task. The first question (question a) in the task does not ask for an exact number but focuses on possible discernment of structure, while the second question (question b) directs attention to the exact number (how many?), to assess students' ways of making use of number structures. In the qualitative analysis, both questions are analysed as one task or phenomenon that the students respond to.

Our interest in this study was primarily directed towards what structures the students discern in the spatial pattern, which was answered qualitatively through carefully analysing notes in the protocols and field notes from the interviews. In this analysis, we took a variation theory approach, which means that we identified in what ways structures were discerned by analysing how parts (the discerned units) and the whole were discerned. For example, a student who is answering ”Four, four and four” but counts all in ones to determine the total number has been interpreted as discerning units of four but they are not related to one another to make a composite whole. A student who answers ”Four, eight, twelve” has on the other hand simultaneously discerned the units as parts of the larger whole and is able to complete the task, thus making use of the number structure s/he has discerned. The interpretations were made based on the students' verbal and gestural reasoning when completing the task. In the results, we illustrate this with figures pointing out the units that the students' express themselves discerning.

Results

In the following, the ways in which the students discern some kind of structure are described. This is related to whether the part-whole relationship is discerned by the students.

The task with the dogs was challenging, as very few students provided a correct number directly and confidently. Because the number of objects to be estimated is relatively large, the students often arrive at an answer by reasoning about subitizable groups that they point out in the pattern. The most common distinctions made were three-groups and four-groups, which is likely influenced by the fact that the pattern is structured as a three-by-four array where the groups of three or four are easily recognizable for the student and often seen as composite units. Special emphasis has been placed on whether students experience these groups as composite or individual units and whether they can discern these related to the whole. The part-whole relationship is regarded as a necessary aspect to discern in order to be able to use number structures for the purpose of determining the number of the objects in the pattern.

In the analysis, four distinct ways of discerning structures in the spatial pattern appear: composite units without relation to the whole, single units in relation to the whole, composite and single units with relation to the whole, and composite units with relation to the whole. These are identified based on how the student discerns units as parts and whole, especially the relationship between them.

Composite units without relation to the whole

First, we observed that some students only discerned parts in the spatial pattern but not a relation to the whole set of 12 objects.

When students discern composite units in the spatial pattern, such as objects in three rows of four or four columns of three in each (see figure 2, left picture), they sometimes cannot discern how these form parts of a larger whole. This can be the student moving or pointing his finger along the columns or rows. For example, one student pointed from left to right along the rows, saying, ”Three fours, then I forgot what it was, so I counted them”, then counting in single units to determine the exact number of dogs. The composite units consisting of three groups of four dogs are then not related to the whole. The student thus expresses the multiplicative relationship in the pattern, but the difficulty lies in seeing how the number of groups is related to the whole. When determining the number of the objects, the student instead discerns only single units and counts them to find the whole.

Another example is a student who discerns units of three by demarcating them within the rows of four objects (see figure 2, right picture). Also, unable to discern how the units of three relate to a composite whole and counts in single units to determine the total number of dogs.

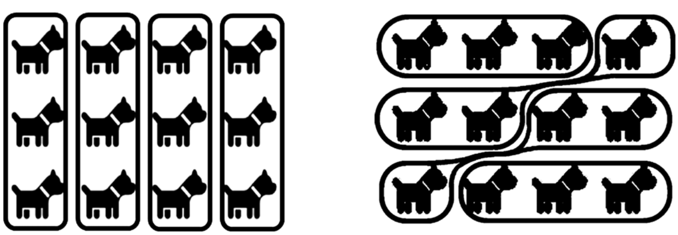

Figure 2. *Students' ways of discerning three-units but with no relation to the whole*

Single units in relation to the whole

Second, students may also discern single units (1) in the pattern, implying a cardinal set (whole) of 12. This is, however not made up of groups as composite units within this whole.

Quite many of the students discern single units. What distinguishes this way of seeing the spatial pattern is that the students count each object in single units, usually following one row or column at a time, to determine the number of dogs. However, no clear structure of parts and whole is used. Students experience the whole consisting of 12 single units but no composite units.

Composite and single units with relation to the whole

Third, we observed several students discerning (at least) two groups in the pattern as composite units greater than 1, forming a whole together. However, the students see the remaining objects in the pattern as single units (1) and counts on to determine the number of objects.

When students discern certain groups in the spatial pattern as composite units greater than 1, forming a whole, some use the ”doubles” structure. This means that they can discern for example, 4 in two rows as 8 (the even parts relate to a whole). However, the student is unable to create additional groups of units larger than 1 to relate to a larger whole than 8, which is why the student continues counting in single units until all objects in the pattern are included in the whole 12, for example pointing at the objects in the rows while saying: ”4 and 8, then 9, 10, 11, 12”. This can be seen as an expression of the student discerning a structure of ”four and four, which is eight” but does not have sufficient experience of how another added four-group forms part of the unfamiliar whole (see figure 3).

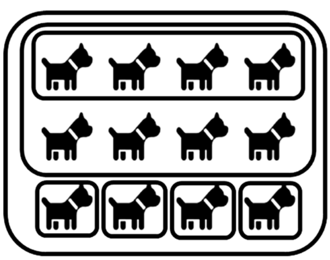

Figure 3. *Student's way of discerning two groups of four as a composite eight but remaining objects as single units within the whole set of twelve*

Different ways of grouping composite units appear in the interviews, followed by one-step counting to determine the exact number of dogs. For example, the student first puts together two or more groups, usually of the same size, and then the units that remain to be counted: ”Here are nine” (covers nine dogs with his hand, then counts) ”9, 10, 11, 12.” In what way the student experiences the nine dogs that are covered with the hand is not clear from the data material, however, the number is greater than a range that is possible to subitize, which is why it is likely that the student discerns some form of structure in the nine dogs (three columns of three in each) while the last column is not expressed as a composite unit but as single units that need to be counted in ones.

Composite units with relation to the whole

Finally, there were students who discerned composite units greater than 1 and related to the whole, simultaneously. The meaning of 12 for these students is a whole within which parts are included in a simultaneously discerned relationship, which helps them determine the number of objects in an effective way.

Some students give a quick and confident answer ”All of them are twelve” and describe the mathematical reasoning that leads to the answer when asked to describe how they arrived at their answer ” ’cause it’s four and four and four”. However, fast correct answers are rarely given in this task. The students who answer quickly and correctly in this way answer the total number already when they receive the first question (”how can you see ...”). The structure of the spatial pattern is simultaneously discerned as a relation between parts and whole – the number is experienced as one object. The student can differentiate parts of the whole and explain how the numbers relate to each other. Most students who answer quickly and confidently can make that differentiation when given a follow-up question. However, they do not need to start by first distinguishing separate parts to see how they form a composite whole.

Many of the students discern units of three or four in the spatial pattern and can also give the answer 12 to the question of how many there are together. They thus discern equal-sized groups forming the whole 12: ”three, three, three and three. Twelve”. As the rows or columns visually form units to focus on, they become prominent as parts of the whole 12. A similar way of reasoning can be seen in students who also discern units of three or four, but in a clear additive structure: ”Four and four makes eight, then four again makes 12” (while pointing to the rows).

Another way of reasoning based on discerned even units involves a simultaneous addition where each part is added to the previous one to form a new whole: ”four, eight, twelve” (see figure 4), while the number grows successively in the student’s reasoning in even steps that include the previously discerned composite units as a part of a new whole. This can be described as the student

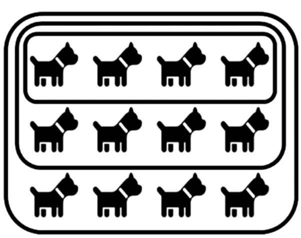

Figure 4. *Students’ way of discerning units within units*

seeing units within units and is able to use this kind of structure to determine the number of dogs.

Discussion

In this paper we set out to investigate in what ways preschool class students discern number structures to determine the number of a spatial pattern (a three-by-four array). The qualitative analysis results in three qualitatively different ways of discerning structures based on composite units, while students who only distinguish single units that are counted in ones (1) do not have the meaning of number structure in the sense of discerning composite units in the spatial pattern. Our study has demonstrated that the discerned composite units are not necessarily related to a cohesive whole at the same time, which hinders the use of the number structures to determine the number of a set. The majority of students turn out *not* to be able to use number structures in a successful way to determine the number of objects, which directs attention to early mathematics education's practices of making the part-whole *relationship* of numbers visible to the students, as units within units. The analysis presented here may give insights into what the teaching should emphasize because of the complex nature of discerning number structure found in this study. In particular, the importance of students being able to attend to and having both the composite units and the whole in their awareness simultaneously.

The number range could play a role in how determining the number of a set can be solved. For example, a lower number range can be memorized or recognized as a spatial pattern representing a certain number, such as dice patterns (Mandler & Shebo, 1982). A larger number range would pose a challenge because it becomes more difficult to memorize certain patterns. However, in Schöner and Benz (2018) studies, students are found to have the same difficulty in discerning parts of a whole for numbers less than ten, as we see in the larger number range up to twelve; that is, students are able to distinguish and relate composite units to a whole but then often count single units to determine the total number. The reason students can discern units but not use the number structures, therefore needs an explanation other than that the number range is too large.

Based on the observations made in the study – that some students have learned to discern units but do not use number structures – it can be concluded that the focus on the part-whole relationship is a necessary aspect, but perhaps not sufficient, for all students to use number structures to determine the number of a set. In the qualitative analysis, the whole emerges as a significant aspect where, in particular the simultaneous discernment of parts (composite units) related to the whole (a structure consisting of composite units) seems to play a role in how the students are able to determine numbers. A didactically important question then arises: whether the teaching focuses on what discerned units are parts of, or in other words, emphasizing the parts, the whole and the relations between them simultaneously. Another aspect to consider in education may be to create the conditions for the student to experience ten as a composite unit because it facilitates determining the number of larger sets, where groups of two, three or four units become a demanding procedure if clear reference points such as five or ten cannot be related to. On the other hand, the pattern in the analyzed task does not provide direct guidance to discern (five or) ten as a composite unit, which may be why very few of the students in this study could distinguish such a structure. One limitation of our study is thereby that only one task has been analyzed. Analyzing additional tasks that include different spatial patterns could provide further insights into the structuring process.

In conclusion, we argue that a dedicated focus on studying and paying attention to students' discerning number structures in didactic research can provide valuable insights for the development of teaching that contributes not only to students' mathematics skills but also to a fundamental and deep understanding of number and how numbers can be structured to see and use relationships within and between numbers successfully.

Acknowledgements

The study was financed by the Swedish Research Council (Grant no. 2020-03712).

References

Baroody, A. J. (1987). *Children's mathematical thinking*. Teachers College.

Baroody, A. & Purpura, D. (2017). Early number and operations: whole numbers. In J. Cai (Ed.), *Compendium research in mathematics education* (pp. 308–354). NCTM.

Benz, C. (2014). Identifying quantities – children's constructions to compose collections from parts or decompose collections into parts. In U. Kortenkamp, B. Brandt, C. Benz, G. Krummheuer, S. Ladel & R. Vogel (Eds.), *Early mathematics learning. Selected papers of the POEM 2012 conference* (pp. 189–203). Springer. doi: 10.1007/978-1-4614-4678-1_12

Björklund, C., Marton, F. & Kullberg, A. (2021). What is to be learnt? Critical aspects of elementary arithmetic skills. *Educational Studies in Mathematics*, 107 (2), 261–284. doi: 10.1007/s10649-021-10045-0

Clements, D., Sarama, J. & MacDonald, B. L. (2019). Subitizing: the neglected quantifier. In A. Norton & M. W. Alibali (Eds.), *Constructing number. Research in mathematics education* (pp. 13–45). Springer. doi: 10.1007/978-3-030-00491-0_2

Ellemor-Collins, D. & Wright, R. B. (2009). Structuring numbers 1 to 20: developing facile addition and subtraction. *Mathematics Education Research Journal*, 21 (2), 50–75. doi: 10.1007/BF03217545

Fuson, K. (1992). Research on whole number addition and subtraction. I D. Grouws (Eds.), *Handbook of research on mathematics teaching and learning* (pp. 243–275). Macmillan.

Hunting, R. P. (2003). Part-whole number knowledge in preschool children. *Journal of Mathematical Behaviour*, 22 (3), 217–235.

Kaufman, E. L., Lord, M. W., Reese, T. W. & Volkmann, J. (1949). The discrimination of visual number. *American Journal of Psychology*, 62, 498–525.

Mandler, G. & Shebo, B. J. (1982). Subitizing: an analysis of its component processes. *Journal of Experimental Psychology: General*, 111 (1), 1–22.

Marton, F. & Booth, S. (1997). *Learning and awareness*. Lawrence Erlbaum.

Mulligan, J., Mitchelmore M. & Stephanou A. (2015). *PASA Response Booklet 2*. Australian Council for Educational Research.

Neuman, D. (1987). *The origin of arithmetic skills*. Acta Universitatis Gothoburgensis.

Paliwal, V. & Baroody, A. J. (2020). Cardinality principle understanding: the role of focusing on the subitizing ability. *ZDM*, 52 (4), 649–661.

Runesson, U. (2006). What is it possible to learn? On variation as a necessary condition for learning. *Scandinavian Journal of Educational Research*, 50 (4), 397–410.

Runesson Kempe, U., Björklund, C. & Kullberg, A. (2022). Single unit counting – an impediment for arithmetic learning. In C. Fernández, S. Llinares, A. Gutiérrez & N. Planas (Eds.), *Proceedings of PME 45* (pp. 363–370). PME.

Schöner, P. & Benz, C. (2018). Visual structuring processes of children when determining the cardinality of sets – the contribution of eye-tracking. In C. Benz, H. Gasteiger, A. S. Steinweg, P. Schöner, H. Vollmuth & J. Zöllner (Eds.), *Early mathematics learning – Selected papers from the POEM Conference 2016* (pp. 123–143). Springer.

Sprenger, P. & Benz, C. (2020). Children's perception of structures when determining cardinality of sets – results of an eye-tracking study with 5-year-old children. *ZDM*, 52 (4), 753–765. doi: 10.1007/s11858-020-01137-x

Steffe, L. P. (1991). Operations that generate quantity. *Learning and individual differences*, 3 (1), 61–82.

Venkat, H., Askew, M., Watson, A. & Mason, J. (2019). Architecture of mathematical structure. *For the Learning of Mathematics*, 39 (1), 13–17.

Wynn, K. (1998). Psychological foundations of number: numerical competence in human infants. *Trends in Cognitive Sciences*, 2 (8), 296–303.

Note

1 The task is inspired by Mulligan, Mitchelmore and Stephanou (2015).

Task frameworks and teacher practice: a comparative analysis

Julia Tsygan and Ola Helenius

We compare three different mathematical task frameworks by applying them to a teacher-described task implementation. By means of a case study method, we juxtapose the frameworks against the teacher's own account. We find that certain crucial facets of the teacher's task implementation remain unaddressed by the frameworks. We posit that this discrepancy arises from the inherently specialized nature of researcher-designed frameworks, which stem from the researchers' theoretical orientations, in contrast to teachers' complex experiences of task implementation. Consequently, we advocate for the development of more comprehensive frameworks to facilitate genuine teacher-researcher collaborations in task design.

Tasks play a central role in mathematics instruction. It is often argued that tasks with a certain richness are vital for fulfilling some general goals and needs in mathematics education (e.g. Krainer, 1993). For this reason, and because teachers often orchestrate student learning using tasks, education designers seeking to improve mathematics teaching and learning often adopt a ”task-centric” design (Boston & Smith, 2011).

Such task-centric design research efforts require analysis of the practicality and effectiveness of rich tasks in regular classrooms, which in turn entails analysis of the opportunities for cognitive challenge present in the task as it is a) formulated by task designers, b) set-up by the teacher and c) implemented by the teacher and students (Tekkumru-Kisa et al., 2020). Likewise, in a reflection on the articles in an edition of ZDM Mathematics Education focusing on task design, Thanheiser (2017) states that ”tasks cannot be considered independently from their enactment”.

The focus on concepts such as intention, alignment, implementation, and enactment, when used in conjunction with the idea of design, draws something of a line between the designer, who is often also a researcher, and the teacher thought of as an implementer of the designer's ideas. The researcher-teacher divide becomes particularly salient when formulated in terms of fidelity

Julia Tsygan, University of Gothenburg
Ola Helenius, University of Gothenburg

(Century & Cassata, 2016). A fidelity focus may indeed promote something of a defect-oriented perspective on the teacher, as either capable or not capable of carrying out the intentions of the designer. It can be argued that such a divide goes quite deep. The work for the researcher can be summarized as focusing on certain theoretically supported and explicitly stated perspectives and providing specific descriptions of what is focused on. The teacher, on the other hand, must attend to many issues at once in the classroom and ”develop implicit theories of action in order to make professional life tolerable [as there are] too many variables to take into account at once” (Eraut, 1994, p. 34).

This brings us to the issue of frameworks developed to analyze tasks. Typically, being developed by researchers, such frameworks should reasonably come with some explicit or implicit foci reflecting the researcher's interest. It is, therefore, interesting to ask what happens when such frameworks are used to analyze the task choice and implementation of an experienced teacher. In this study we consider the case of a task chosen, modified, and implemented by a teacher, and apply three different researcher developed frameworks to it. Using case studies for ”theory testing” (Thomas, 2011, p. 516) has been done previously. For example, Phipps and Merisotis, (2000) used the case study method to investigate the alignment between a framework that describes quality in distance education with the actual practices as perceived by university leaders, teachers, and students. We would like to follow Phipps and Merisotis' approach, albeit in a much smaller scale, to use a case of a teacher implementing a task in order to test common frameworks of task implementation.

The three frameworks chosen for this investigation were selected based on being a) similar in concerning themselves with cognitively challenging tasks, b) different in how fine-grained analysis they offer, and c) stemming from three different academic contexts. Among the three frameworks, two have a quite analytical focus (COACTIV and Stein's) and one has an explicit design perspective (CMR). This means that the combined selection of frameworks has the potential to capture different aspects of task characteristics and implementation when applied to a specific case of task implementation.

Could it be that an experienced teacher's use and perspective on rich tasks may contain a selection of issues that established task analysis and task design frameworks will not capture? This is what we will examine in this paper by applying the three task frameworks to a teacher's task formulation and described implementation. The research question we ask is: *What aspects of a teacher's task implementation can a selection of these three task frameworks capture?*

Description of the frameworks

Among multiple frameworks available in the research literature, three commonly used frameworks will be considered: Stein's framework for analysis of mathematical tasks (Stein et al., 1996), *Cognitive activation in the* *mathematics classroom* (COACTIV) (Neubrand et al., 2013), and Lithner's "*Creative mathematically founded reasoning* framework for task design" (CMR) (Lithner, 2017). In this section, we summarize the main ideas and applications of the frameworks and leave the details of these frameworks for the Results section where they are presented in relation to the teacher's task formulation and description of implementation.

In Stein's (1996) framework, tasks can be considered in terms of their inherent features (qualities of the task formulation itself) and cognitive demands (capacities students must call on to successfully solve the task). The task features and cognitive demands are brought to the students through the teacher's framing of the implementation of the task, which influences but does not determine student enactment of the task. Successful implementation of rich tasks depends here on several teacher-controlled variables of which building on prior learning, scaffolding, appropriate timing, modelling of high-level performance, and sustained pressure for explanations or justifications, are seen as especially important (Henningsen & Stein, 1997). Since its publication in 1996, Stein's framework has been extensively cited in research on teachers and students as agents in the use of tasks.

Within the COACTIV framework, the task is analyzed according to ten different categories thought to elicit higher levels of cognitive activation in learners that engage with the tasks, with each category evaluated on a numerical scale (Neubrand, 2013). The COACTIV classification framework is focused on the task as intended or formulated (rather than the framing for its implementation) and has primarily been used to analyze written formulations of tasks (most notably in Neubrand, 2013). The classification framework does not explicitly address classroom implementation. Instead, the related COACTIV Quality instruction model (Kunter & Voss, 2013), developed for and validated by analysis of teaching in lessons, may help with the analysis of implementation of tasks. This model conceptualizes quality of instruction as being based on cognitive activation, classroom management, and individual learning support quite generally, each consisting of multiple smaller variables.

Lithner's *Creative mathematical reasoning* (CMR) framework (Lithner, 2008) defines CMR as reasoning that is novel, plausible to the reasoner and underpinned by mathematically sound arguments. The CMR framework has been widely applied in analyzing tasks in textbooks and assessments. Lithner's subsequent work in 2017 further developed the CMR framework for task design, advocating for tasks that challenge students' conceptual, creative, and justificatory skills. Here, effective implementation hinges on teachers resisting the urge to excessively aid students, which would otherwise diminish the challenges inherent in the task.

Collectively, Stein's framework, COACTIV, and CMR offer a multidimensional perspective on task analysis and implementation.

Method

Our method can be characterized according to the case study typology provided by Thomas (2011). The *object* of study is the task analysis frameworks' ability to explicate aspects of task characteristics and task implementation. For this, we chose the three frameworks presented above which we see as useful to study as they are relatively well used by both their developers and other researchers.

The *subject* of study is a teacher's implementation of a rich task. The teacher was selected as a ”local knowledge case” (Thomas, 2011), being a colleague of the first author, who is also a teacher at the same school. The proximity and professional connection enable the first author to ascertain that the teacher has a very well-thought-out view of tasks as well as their implementation in her teaching context. This means that the chosen subject presents a meaningful challenge for the frameworks and could be used to evaluate them in a theory testing approach.

We asked our case teacher to provide both a task and an explanation of why she selected it. The teacher emailed us a description of her task formulation and her framing of the implementation of it with her 6th- and 7th-grade students. We conducted our analysis in two distinct stages. Initially, we applied each of our selected frameworks to the task formulation. Following this, we employed inductive thematic analysis at the semantic level (Braun & Clarke, 2006) to identify recurring themes within the teacher's account of the task's implementation. The first author reviewed the text multiple times, identifying raw data themes such as words and phrases pertaining to implementation. These were then organized into overarching themes. We then shared the analysis with the teacher for a member-check to ensure its credibility. Finally, we compared the themes with the criteria from the three frameworks.

Results

In this section, we first analyze the task as it was formulated by the teacher. Then we present and analyze the teacher's description of the implementation of the task.

Application of frameworks to the task formulation

The task, presented in figure 1, is an example of a ”puzzle task” (called so by the teacher with her students) used by the teacher.

The teacher reported that she found the original task online, and then rephrased and expanded the task significantly. When asked, she described her goals:

> It depends on the topic, but since they have to look really hard at something I often try to make it something many kids miss or something usually taught by just telling it to kids, which they don't digest or retain. [...]

> Other than that my general goal is helping them learn to think and speak precisely, and to start to use deductive reasoning to show a conclusion, a baby step to proof.

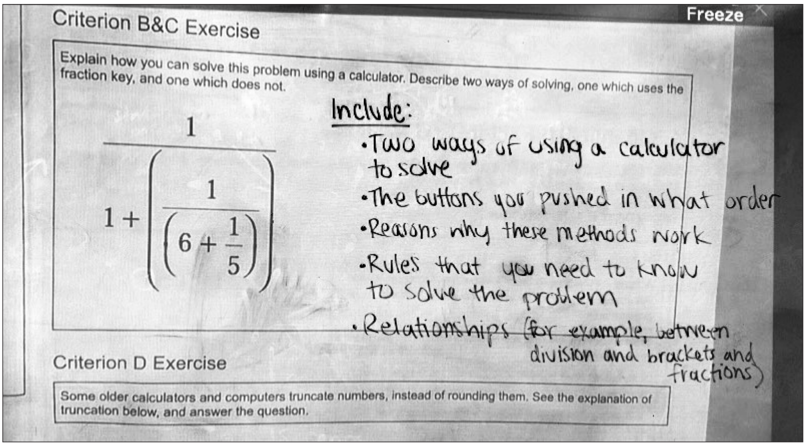

Figure 1. *Photo of the problem presented to students in a lesson*

The teacher's goals for using such tasks thus seem both content-specific and intended to further the gradual development of complex reasoning and communication skills.

Using Stein's framework, the task features include the requirement to produce multiple solutions and justifications. The use of a calculator could also be interpreted as a form of representation of the fraction in the problem. In terms of cognitive demands, a successful solution requires a choice of procedures which in turn requires conceptual understanding of the grouping of numbers by brackets and divisions.

The scores for COACTIV criteria can be found in table 1. The scale ranges from 0 to 3 for each criterion. An asterisk (*) or two (**) indicates that it is difficult or very difficult, respectively, to determine the score for that criterion, according to the authors' opinion. The task formulation is thus likely to receive high scores on some COACTIV categories and low scores on others.

Under the CMR framework, the given task lacks a clear solution process and therefore requires some degree of creativity to be solved. It requires conceptual understanding of the relationship between fraction bars and brackets used on a calculator, and explicitly asks for an explanation based on mathematical reasoning, thus satisfying the CMR criteria.

From the summary in table 1, it is evident that the task formulation meets the different frameworks' characterisation of rich tasks that elicit higher-level mathematical thinking.

Table 1. *Framework elements found in the teacher's task description*

COACTIV		Stein		CMR	
Category	*Score*	*Task features*		*Characteristics*	
Topic area (1–4)	1	Multiple solutions	✓	Creativity	✓
Curricular knowledge level (1–3)	3*	Multiple representations	✓	Conceptual challenge	✓
Type of mathematical activity (1–3)	2*	Demand for justification	✓	Demand for justification	✓
Extra-mathematical modelling (0–3)	0	*Cognitive demands*		Reasoning	✓
Inner-mathematical modelling (0–3)	3*	Procedures with conceptual connections	✓		
Basic concepts (0–3)	3				
Processing of mathematical texts (0–3)	0				
Argumentation (0–3)	2				
Direction of task solution (1–2)	1**				
Number of solution paths required (0–2)	2				

The task implementation by the teacher

The teacher's description of her implementation of the task is presented below. The underlined text is coded to show the emergent overarching themes:

(1) – *timing* considerations

(2) – *feedback or scaffolding*

(3) – *push for justification* or explanation

(4) – *modelling* appropriate engagement with the task

(5) – *curriculum context, prior learning*

(6) – attention to *assessment demands*

(7) – *sequencing* of multiple tasks over time

Many words and phrases in the teacher's account encompassed multiple themes present in this analysis. Notably, data coded as Sequencing also connects with Curriculum as well as Scaffolding. The number code thus reflects only the main focus of each segment.

> I tell the students that once per term (1) they should choose two of the B & C puzzles (6)from the curriculum we have completed so far (5) (which they have worked on in class for roughly 30–45 minutes per task) (1) and develop them to turn in as part of their summative portfolio (6).
>
> In the lesson, I put the puzzle on the board and read it. I ask questions to ensure the students understand what is being asked in the puzzle. I may give simplified

> examples to ensure students understand what is being asked, but I do not tell the students how to start. (6) Then I alternate between a 6 minute timer for working alone, and a 4 minute timer for discussion, and allow the students to attempt to find the pattern/solve the puzzle in roughly 15–20 minutes. (1) I check students' answers individually, and try to ensure everyone has an answer. If students have the wrong answer, I tell them to keep working on the answer, maybe with a small hint (2), and if they have the right answer I ask them to write their explanation/justification (3).
>
> I model an explanation/justification at the start of the year so that students know what this means. Usually, I solve a simple pictorial algebra problem, I solve it, and I write the explanation/justification of my solution, explaining why certain parts (rules I used to solve the problem, why those rules work, relationships between operations or steps, definitions of mathematical terms, etc) are necessary to justify my answer [...] I highlight key words in the explanation/justification which are common in good mathematical justifications, and list good features of them, for example, accounting for all types of numbers (would a rule work with negatives? 0? 1?). (4)
>
> Back to the lesson itself, when students have an answer I ask them to start the explanation (3).During this process, I may point out good features of explanations when we are working on formative versions of the task. I usually specifically invent an explanation that reads like a story and read it in a silly way to show that a justification is not a story of your journey through a problem (4). I give regular feedback during the explanation writing process to let students know where they are missing things. For example, you say you did x, (2) but you give no reason for doing x. Why did you have to do that? (3)
>
> In the first few formative versions of the assignment, I write some key words or prompts on the board to help students decide which of their thoughts would be a good thing to include in their explanation (7). See the example attached. [figure 1]
>
> [....] I refer to the rubric for more detail, which we have discussed in class. I include the rubric also. (6)
>
> Usually before the due date of the summative we have one or two explanation workshop days where students work on these (7) and I give regular feedback to let them know to what extent they are on track, and if not, why not. (2)

During the member check, the teacher agreed with the results and reported that, in her opinion, modelling and feedback, together with ”Socratic questioning” (that could be conceived as a type of feedback, authors' note), are the most important elements.

Application of frameworks to the analysis of task implementation

A summary of the extent to which the frameworks capture the themes in the teacher's description is presented in table 2 and explained below.

Students' work on solving the task is organised in clearly defined time intervals. Of the three frameworks considered here, only Stein explicitly considers appropriate timing as crucial to successful task implementation.

Scaffolding or supporting student problem-solving is mentioned in all three frameworks, although COACTIV is less explicit than Stein and CMR about what such support looks like. Likewise, all three frameworks agree with the teacher that the demand for justification is an important element of the task.

Modelling features prominently in the teacher's description of her teaching, as well as in her statement (during member check) that this is among the more important elements in her implementation of the task. Stein's is the only framework that considers the modelling as an important factor affecting implementation.

Curricular context and building on prior learning are mentioned explicitly in both COACTIV and Stein, although COACTIV does not consider this a factor in implementation but rather a fact of the task formulation itself.

Some important aspects of the implementation of this task were not captured by any of the frameworks: the prominent role of formative and summative assessment and the positioning of the task among other tasks of a similar kind.

Table 2. *Themes from the teacher's description captured by the frameworks*

Themes in the teacher's description	Stein	COACTIV	CMR
Timing	✓		
Feedback or scaffolding	✓	✓	✓
Push for justification	✓	✓	✓
Modelling engagement	✓		
Curriculum context, prior learning	✓	✓	
Assessment demands			
Sequencing			

Discussion

We examined the application of three frameworks to a quite advanced and well-thought-out task, chosen, modified, and implemented by an experienced teacher. While the task formulation fit well with the frameworks' characterization of task design, only Stein's framework was effective in elucidating critical aspects of task implementation, aligning closely with the teacher's description. While COACTIV offered a more granular understanding of potential opportunities embedded in the task formulation, some categories proved challenging to assess, and the relevance to the teacher's specific implementation was limited. The task formulation also aligns with CMR requisites, yet the CMR framework appears comparatively constrained when juxtaposed with the rich, multifaceted perspective provided by the teacher.

The question arises: why does CMR, which is the sole design framework in our selection, capture so few of the crucial facets of implementation identified by our teacher? In light of research indicating the pivotal role of teachers in facilitating students' enactment of tasks (Henningsen, 1997), it is reasonable to anticipate that research on CMR task design will develop toward formulating more complex design principles to support teachers in CMR task implementation. Indeed, such research efforts appear to be in progress, as evidenced in Sidenvall et al. (2022).

The teacher's account of her implementation of the task highlighted several considerations that were not explicitly addressed in either framework. These included a focus on assessment requirements, the task's relationship to preceding tasks that differed in content but shared similar cognitive demands, timing that extended over multiple lessons, and a deliberate shift from formative assessment, involving extensive modelling and feedback, to more independently conducted summative assessment at the term's end. Overall, the teacher's implementation considers more facets of teaching and learning, than did either of the frameworks individually or collectively. This aligns with Eraut's (1994) observation that teachers often navigate numerous complexities in their lessons. Generally, it may be the case that certain behavioural routines that teachers find useful in light of their many priorities may not fit well with the intended task implementation. In the model of implementation of cognitively demanding tasks presented by Tekkumru-Kisa et al. (2020), both the teacher's set-up of the task as well as the actual classroom work with the task may inflict decline in cognitive demand. Therefore, frameworks that encompass a broader range of factors, particularly those aligned with teachers' priorities such as assessment demands and the relation to long-term skill development, are more likely to support task implementation where cognitive demand is maintained.

This view fits well with Thanheiser's (2017) emphasis on the importance of evaluating tasks in relation to domain-specific versus domain-transcendent objectives. In this terminology, the task depicted in figure 1 addresses both domain-specific goals, such as facilitating a more profound grasp of concepts that might otherwise easily be forgotten, and domain-transcendent goals, such as the cultivation of precise communication and reasoning skills. The frameworks discussed in this article appear not to account for these extended objectives, nor do they address how a task can be used to serve both purposes.

Ultimately, the frameworks were applied to the teacher's account of the task's implementation rather than the implementation itself, live in the classroom. It is conceivable that using filmed observations during implementation could have unveiled additional pertinent factors that the frameworks may have addressed to varying degrees. Subsequent research should incorporate direct observations of both the teacher's implementation and the students' execution of the task.

Conclusion

In this case study, we are grateful to a teacher who generously allowed us insight into her use of a challenging task, the implementation of which turned out to be more complex than three common frameworks could fully capture. The teacher employs the task for multiple purposes: to improve understanding of the curriculum, help students gradually develop complex skills, and to meet assessment requirements. She strategically places the task within a series of similar activities spanning many weeks, choosing the right moment based on student needs and preferences. This study underscores the need for research frameworks to evolve and encompass a broader range of aspects that are pertinent to teachers. Developing such comprehensive frameworks not only benefits our understanding of teaching practices, but also has the potential to enhance our ability to assess the practicality and effectiveness of tasks within design research and implementation studies.

References

Boston, M. D. & Smith, M. S. (2011). A ”task-centric approach” to professional development: enhancing and sustaining mathematics teachers’ ability to implement cognitively challenging mathematical tasks. *ZDM*, 43, 965–977. doi: 10.1007/s11858-011-0353-2

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3 (2), 77–101. doi: 10.1191/1478088706qp063oa

Century, J. & Cassata, A. (2016). Implementation research: finding common ground on what, how, why, where, and who. *Review of Research in Education*, 40 (1), 169–215. doi: 10.3102/0091732X16665332

Eraut, M. (1994). *Developing professional knowledge and competence*. Psychology Press. doi: 10.4324/9780203486016

Henningsen, M. & Stein, M. K. (1997). Mathematical tasks and student cognition: classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, 28 (5), 524–549. https://doi.org/bh4q7f

Krainer, K. (1993). Powerful tasks: a contribution to a high level of acting and reflecting in mathematics instruction. *Educational Studies in Mathematics*, 24 (1), 65–93. https://doi.org/djkhfw

Kunter, M. & Voss, T. (2013). The model of instructional quality in COACTIV: a multicriteria analysis. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers: results from the COACTIV project* (pp. 97–124). Springer.

Lithner, J. (2008). A research framework for creative and imitative reasoning. *Educational Studies in Mathematics*, 67, 255–276. https://doi.org/cvrffd

Lithner, J. (2017). Principles for designing mathematical tasks that enhance imitative and creative reasoning. *ZDM*, 49 (6), 937–949. https://doi.org/f97rrp

Neubrand, M., Jordan, A., Krauss, S., Blum, W. & Löwen, K. (2013). Task analysis in COACTIV: examining the potential for cognitive activation in German mathematics classrooms. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers: results from the COACTIV project* (pp. 125–144). Springer.

Stein, M. K., Grover, B. W. & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: an analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal*, 33 (2), 455–488. https://doi.org/fmrhsx

Phipps, R. & Merisotis, J. (2000). *Quality on the line: benchmarks for success in internet-based distance education*. Institute for Higher Education Policy. https://eric.ed.gov/?id=ed444407

Sidenvall, J., Granberg, C., Lithner, J. & Palmberg, B. (2022). Supporting teachers in supporting students' mathematical problem solving. *International Journal of Mathematical Education in Science and Technology*, 1–21. https://doi.org/grj8mn

Tekkumru-Kisa, M., Stein, M. K. & Doyle, W. (2020). Theory and research on tasks revisited: task as a context for students' thinking in the era of ambitious reforms in mathematics and science. *Educational Researcher*, 49 (8), 606–617. https://doi.org/ghjspk

Thanheiser, E. (2017). Commentary on mathematical tasks and the student: coherence and connectedness of mathematics, cycles of task design, and context of implementation. *ZDM*, 49 (6), 965–969. https://doi.org/mcjx

Thomas, G. (2011). A typology for the case study in social science following a review of definition, discourse, and structure. *Qualitative Inquiry*, 17 (6), 511–521. https://doi.org/d928rt

Practising feedback: use of role-play in online mathematics teacher education

Kristofer Sidenvall

This paper discusses the use of role-playing in online mathematics teacher education for preparing pre-service secondary mathematics teachers to use formative feedback. Findings from video-recorded role-plays and participant reflections highlight how the context presented in the online role-plays can enable or restrict different types of feedback. The study also reveals that the online environment supports rehearsal for giving feedback but limits participants' ability to express feedback. The paper offers suggestions for enhancing the potential of online role-playing for feedback rehearsal.

In mathematics teacher education (MTE), a practice-based approach emphasises regular practical experience, immersing pre-service teachers in real teaching situations (Ball & Cohen, 1999). This approach aims to enhance teaching skills through hands-on experiences during field placements and, within university courses, approximated mathematics teaching scenarios like role-playing, video observation, and co-teaching in controlled settings (Grossman et al., 2009). Although previous research has verified the benefits of a practice-based approach to MTE when employed in campus-based settings, a review of research on practice-based teacher education (Matsumoto-Royo & Ramírez-Montoya, 2021) revealed no examples of studies in online settings.

MTE should include preparing pre-service teachers to provide feedback on pupils' thinking. Assessment and feedback on pupil understanding are crucial for pupils' learning, and effective formative feedback significantly enhances pupil achievement (Anthony & Walshaw, 2009; Hattie & Timperley, 2007). Despite its potential, reports reveal shortcomings in the quality of feedback given by teachers to pupils (e.g. Hirsh & Lindberg, 2015; Stovner & Klette, 2022). Consequently, during their field placements, pre-service teachers may not encounter the kinds of feedback practices that research advocates.

The focus of this paper is the implementation of a practice-based approach in online MTE where pre-service teachers practise giving different types of

Kristofer Sidenvall
Dalarna University

feedback in approximated mathematics education settings. One effective method for preparing pre-service teachers for demanding teaching situations, such as giving formative feedback in the moment, is to use role-playing (Lajoie, 2018). Based on this, the question to be answered is: *What does role-play as an approximation of teaching offer online MTE in terms of rehearsing for feedback?* This will be answered by analysing recorded online role-plays and the discussions about participants' experiences of the role-plays.

Previous research

In this section, previous research on practice-based MTE, online education, teaching approximations such as role-playing, and elements of a formative feedback process in mathematics teaching will be presented.

Practice-based MTE involves regular rehearsal for pre-service teachers in real or made-up teaching situations, aiming to improve their teaching skills through hands-on experiences in field placements but also during university courses in approximated mathematics teaching scenarios (Ball & Cohen, 1999; Grossman et al., 2009). There are various examples of exercises that approximate teaching practices such as lesson planning, role-playing, co-teaching, and reflection writing (McDonald et al., 2013). One goal of approximations is preparing pre-service teachers to make better decisions when faced with teaching challenges (Lampert et al., 2013). Here, role-playing has been highlighted as a powerful approximation for preparing pre-service teachers to act in the moment (Lajoie, 2018). The relevance of role-play for such preparation is motivated by its authenticity in relation to the real teaching situation (Howell & Mikeska, 2021). Owing to the nature of approximations such as role-play, teaching situations can be simulated to prepare teachers for mathematics teaching in safe, controlled, and scaffolded environments with the support of a mathematics teacher educator (Grossman et al., 2009).

As more students enrol in online MTE programmes (Dyment & Downing, 2020), questions arise about adapting practice-based MTE to virtual settings. Research has shown the benefits of this approach in campus-based contexts for both pre-service and novice in-service teachers (Kazemi et al., 2015; McDonald et al., 2013) but there is limited research on adapting practice-based approaches to online contexts (Matsumoto-Royo & Ramírez-Montoya, 2021). Nonetheless, online mathematics education faces challenges related to instruction, interaction, and collaboration (Radmehr & Goodchild, 2022).

Feedback on pupils' learning is one of the teaching practices that are central to a mathematics teacher's work (McDonald et al., 2013). Feedback can serve both summative and formative purposes, the latter being emphasised as particularly effective for pupils' learning (Hattie & Timperley, 2007; Ryve et al., 2015). Formative feedback can be defined in different ways in various contexts.

In educational contexts, formative feedback refers to information about pupils' performance and learning that aims to enhance achievement and bridge the gap between actual and desired learning (Hattie & Timperley, 2007).

Because a teacher provides feedback to a pupil in the classroom it involves some form of communication. For the pupil's learning, it is important that the conversation focuses on the pupil and their ideas rather than on the teacher's opinions (Hufferd-Ackles et al., 2004). Similar to Ruiz-Primo and Li's (2013) perspective, in this paper the feedback concept also encompasses communication during which the teacher elicits the pupil's thinking as a step towards providing feedback. This expansion of the concept of feedback means that both the acting teacher and the acting pupil become important for the analysis. A deeper explanation of what the feedback process covers in this paper will follow in the theory and methodology section.

Research has shown that formative feedback is something teachers spend much time on (Björklund Boistrup, 2022; Stovner & Klette, 2022), and the appreciation of the potential of formative feedback is well known (Hirsh & Lindberg, 2015). Despite this, research finds deficiencies in the quality of the feedback given (e.g. Hattie & Timperley, 2007; Hirsh & Lindberg, 2015; Stovner & Klette, 2022). Given this background, it is essential to examine feedback rehearsal when incorporating role-playing into online education within the context of MTE.

Theory

The design of the study is informed by the experimental cyclic learning process theory used by Kazemi et al. (2015). It involves four phases: (1) introduction to a teaching practice, (2) meeting for collaborative preparation, (3) practice enactment with pupils in a mathematics classroom, and (4) reflective analysis, which brings experiences into the next cycle. This paper reports on data from the enactment and analysis phase. Modifications were made to enable the enactment phase fit the current study's context. Instead of enacting the feedback practice in traditional classroom interactions in a school, participants engage in approximated teaching scenarios in an online environment where one of the participants acting as a pupil replaces the authentic pupil. In this way, a recurring enactment by practising to provide feedback to pupils becomes enabled in online MTE even without access to a real classroom.

The feedback process

Previous studies have delineated various components of formative feedback, and relevant parts of these theoretical contributions have been selected to lend support to this study. The four categories – task focus, process focus, self-regulation focus, and self-focus, based on Hattie and Timperley (2007), capture some characteristics of expressed feedback, with process and self-regulation-focused feedback being the most effective, and self-focused feedback is less efficient. However, combining self-focused feedback with task or process feedback can mitigate the negative effects of feedback about the individual, e.g. repeated praise (Hattie & Timperley, 2007). Feedback can be communicated using various expression forms, encompassing verbal communication, written text, laboratory materials, images, symbols, body language, and more (Björklund Boistrup, 2022). These expression forms are important in shaping the meaning of a discourse (Van Leeuwen, 2005). One expression form of feedback involves teachers being silent. Despite this being a practice that leads to improved quality and quantity of pupil responses, teachers rarely allow a pupil more than three seconds to reply (Björklund Boistrup, 2022; Rowe, 2003).

The communication between teacher and pupil during feedback processes varies and involves the pupil to differing extents. This study considers teacher-pupil communication to be an integral part of the feedback process. Consequently, the framework integrates two levels of communication characteristics, drawing inspiration from Hufferd-Ackles et al. (2004) who advocate a transition from a teacher-centred to a more pupil-focused approach, promoting increased pupil engagement. Unlike Hufferd-Ackles et al., only levels 0 and 1 are used in this study, because levels 2 and 3 assume several pupils are involved. Levels 0 and 1 are defined later in this paper.

Methodology

Context and participants

The participants were students in their eighth semester in a Swedish online mathematics secondary teacher education. At the time of the study, these students were in a course that brought together pre-service teachers from various subjects. All three students specialising in mathematics consented to participate. The two meetings for data collection, hereafter called activities 1 and 2, were separate from participation in the university course. Although the researcher taught other students in the university course, he was not involved in the three participants' studies. This helped to reduce power imbalances (Hayes et al., 2008), and to ensure that the study was not compromised by his having dual roles.

Design of activities

To generate data pertaining to the research question, during activity 1, role-plays in which participants assumed the roles of either a teacher or a pupil were conducted. The role-play session was preceded by an introduction to the subject of feedback in university course materials and participants shared their ideas about feedback before they enacted their ideas in the role-plays. The role-plays were based on tasks from various areas of mathematics to demonstrate the difficulties pupils commonly experience in mathematics education (e.g. Di Lonardo Burr et al., 2020). The purpose of the choice of tasks was not to challenge the participants' mathematical knowledge but to form a basis for mathematics teaching feedback. Every role-play started with a moment of individual reflection on a given mathematical task and a fictional pupil's related incorrect written solution. Then, the acting teacher asked questions to the acting pupil to understand the pupil's thinking and then responded accordingly. It is worth noting that the role-play was not about acting as an actor, but expressing what was found appropriate for the task. Based on the recommendations of Shaughnessy and Boerst (2018), the acting pupil was told to restrict their responses to explicitly answering the teacher's question, to challenge the teacher's elicitation skills. During the role-plays, the researcher coordinated, gave instructions, selected tasks, distributed roles, and allocated time between different exercises. During the analysis phase – activity 2, an online follow-up discussion a week later – the participants' impressions about the possibilities and limitations of using role-play for feedback learning were collected.

Data analysis

Video recordings of activities 1 and 2 were automatically transcribed with transcription software and then manually corrected. Those parts particularly relevant to the research question were edited and cleaned, because a verbatim transcript was not considered useful for this paper. The data analysis aims to comprehensively address the research question through two different approaches. The first, a thematic analysis (Braun & Clarke, 2022) with a deductive orientation was employed to describe feedback expressed in the role-plays. For the deductive analysis, NVivo was used to pair text to seven preexisting codes linked to the feedback process. Classification criteria for five of these codes are described in table 1. Alongside these five codes, the teacher-pupil communication was coded at either level 0 or 1. At level 0, the teacher is telling and showing. The pupil is not expected to elaborate on the frequent and short questions, only reply with the expected answer. The pupil's answers are not followed up by probing follow-up questions. At level 1, the focus of the teacher's questions has moved towards the pupil's thinking. The teacher asks more probing questions, and the pupil becomes more involved and has an impact on their own learning.

The second approach captures the specifics of role-playing and its implementation in an online environment. An inductively oriented thematic analysis was applied, discovering content that contributed to a deeper understanding of how role-playing in online MTE can be an opportunity to practise feedback.

Table 1. *Thematic framework for feedback process analysis*

Codes	Criteria, characteristics	Examples
Task focus	Focuses on something specific in the task at hand. Corrects, without involving processes. Provides solutions. Gives advice that elicits a particular answer or strategy.	*Now, subtract 93 and 47; so first mark 93 on the number line.*
Process focus	Pays attention to process or strategy beyond the task itself. Provides generalisable hints, even if they are targeted at the task. "Sketch a graph." Asks for meanings and definitions of different concepts. Corrects, by involving processes.	*Why is it like that when you add fractions?*
Self-regulation focus	Provides feedback aimed at the pupil, developing their ability to control their own learning. Supports the pupil to review and re-count solutions.	*Is the answer reasonable? Test your solutions.*
Self-focus	Feedback only relates to the individual.	*Good job!*
Expression form	Besides verbally, feedback can be expressed by using or articulating the use of or need for written text, symbols, pictures, images, laboratory materials, body language, or if the acting teacher stays quiet for over three seconds while waiting for the pupil to reply to a question.	

Note. The framework is based on Björklund Boistrup (2022), Hattie and Timperley (2007) and Rowe (2003). The codes for teacher-pupil communication (Hufferd-Ackles et al., 2004) are not displayed in the table.

During an open coding process, each line and quotation was labelled pertaining to feedback, the online environment, and the role-play content of the transcripts.

Through the process of searching, reviewing, and refining patterns among the findings from the deductively and the inductively oriented thematic analysis, three themes emerged that captured meaningful data relevant to the research question: *Feedback independent of teaching-pupil information, context-specific challenges* and *role-play-specific reflections*.

Results

In this section, results salient to each theme are presented.

Feedback independent of teaching-pupil information

The analysis of the two recorded activities showed that throughout the various role-plays, no feedback was expressed regarding self-regulation or the self as a person. This is probably connected with participant reflections on the lack of context such as knowledge of the pupil and specific goals. Namely, the instructions for each role-play did not provide information about goals or the fictional pupil's prior knowledge or social situation. Such information is what this theme refers to as teaching-pupil information. However, at the same time, role-playing can lead to the use of feedback focusing on both the task and the process. To the task with the incorrect answer, $93-47=54$, the acting teacher suggests the acting pupil use a number line.

Teacher: If you draw a number line? Can you calculate the same task with the number line? Could you explain to me step by step how to do it?

Pupil: What do you want me to count then? The difference between 3 and 7?

Teacher: No, 93 and 47; so mark 93 on the number line.

Pupil: Yes.

Teacher: And then from 93 you need to subtract 47, but you don't have to do it in one step. Ehm ... without ... How do you think you can more easily subtract it on the number line?

In the initial part of the excerpt, the teacher's feedback is task oriented, but it transitions to addressing the pupil's understanding of subtraction on the process level in the final sentence. Despite the absence of knowledge about the pupil, the conversations within the role-plays can prioritise the pupils' thinking. In the preceding passage, the communication occurs at level 0 until the final sentence, at which point the teacher shifts to level 1. A more distinct illustration of communication at level 1 will be evident in the next role-play excerpt. To the incorrect solution, $\frac{2}{7} \times \frac{3}{7} = \frac{6}{7}$, the acting teacher probes the acting pupil's approach by asking "What are you going to do, in the task?" It becomes clear that the pupil understands the need to multiply fractions but incorrectly believes that the denominators should remain the same when multiplying fractions with equal denominators. The teacher continues.

Teacher: Is it the same rule when adding fractions?

[Silence, 8 sec.]

Pupil: Yes, I think so.

Teacher: Why is it like that when you add fractions?

[Silence, 5 sec.]

Pupil: When we add the two?

Teacher: Yes, if there had been an addition sign instead of a multiplication sign?

Pupil: Since we have the same denominator then it equals ... [short silence] ... five times ... 5 divided by 7.

Communication thus far is at level 1 and focused on processes, featuring an instance where the teacher employs silence. The teacher continues the feedback process by prompting the pupil to contemplate another multiplication, hoping this will lead to insights that assist in solving the initial problem. In contrast to the previous excerpt, the teacher now transitions to communication level 0 but maintains a process focus with the feedback. This feedback may also serve a self-regulatory purpose if the pupil can apply the strategy to solve similar, simpler problems.

Context-specific obstacles

The teacher's introduction of a number line in the first excerpt of the previous theme suggests an attempt to incorporate non-verbal forms of feedback. Later in that role-play, the pupil displays their attempt at a drawn number line in the webcam. Here, the limitations of the digital online environment are illustrated as efforts to transition between various forms of expression are impeded by contextual constraints. During the analysis phase, participants also acknowledged the limitations posed by the online context, especially when it came to mathematics instruction. The online environment was perceived as a safer space for role-play compared to in-person meetings, although gestures or facial expressions were difficult to interpret.

Role-play-specific reflections

The participants viewed the role-plays as a valuable opportunity to practise teaching, ask questions, and provide helpful feedback that goes beyond simply describing rules or solutions. Participants also found value in exploring different forms of feedback and becoming more aware of their own feedback style. Among the participants, previous experiences of approximations of teaching were limited to reflexive writing and discussions about pupils' written solutions. Additionally, participants suggested more time for discussion after every role-play and an actively participating mathematics teacher educator.

Discussion

Initially, the question was asked: What does role-play as an approximation of teaching offer online MTE in terms of rehearsing for feedback? Communication during the feedback process was observed at levels 0 and 1, with instances of participants transitioning between the two. In this way, the role-plays illuminate the communication characteristics as described by Hufferd-Ackles et al. (2004) in the feedback process. Shifting between communication levels may result from the focus on feedback in the role-plays, because participants viewed the role-plays as opportunities to test various feedback types. Participants also talked about limitations in providing feedback due to the limited knowledge of the fictional pupils and the scarcity of clear goals for the mathematics teaching situation presented in each role-play. These limitations may have contributed to the almost total absence of self-regulation and self-focused feedback. In contrast, the role-plays yielded repeated feedback on both the processes and tasks. Creating informative contexts around the exercises and pupils' situations may enhance the authenticity of role-plays, as Howell and Mikeska (2021) emphasise, which might improve the conditions for providing diverse and well-grounded feedback in these scenarios.

In the online environment, limitations became apparent when there was a need to use symbols spontaneously, and alternative means of expression beyond verbal communication were sought. These obstacles may arise from the specific demands of online mathematics teaching (Radmehr & Goodchild, 2022). Participants attempted diverse expression forms during the feedback process, including silence and visual aids like number lines, and expressed difficulties in interpreting body language. Although these modes of expression align with the ideas of Björklund Boistrup (2022), Rowe (2003), and Van Leeuwen (2005) regarding the feedback process, they are constrained within the virtual environment. However, these challenges underscore the importance of preparing the exercises and exploring alternative digital resources, such as synchronous drawing tools, to facilitate diverse forms of expression.

Lampert et al. (2013) and Lajoie (2018) emphasise the importance of practising challenging parts of teaching and preparing to act in the moment. These opportunities can be seen in online role-plays because the acting teacher has to decide how to respond based on the information they have just received from the acting pupil at that moment. In their reflections, participants also placed great value on being able to practise acting as a teacher, responding to pupils, and testing different types of feedback. The willingness to test and experiment with feedback can be attributed to the controlled and scaffolded environment that approximations provide (Grossman et al., 2009). Participants also specifically expressed feeling secure in the online environment. In contrast to the participants' prior experiences of approximations, which according to Howell and Mikeska (2021) can be considered less authentic for in-the-moment feedback, the role-plays offer the opportunity to enact feedback practices in addition to investigating feedback practices.

It is essential to acknowledge the potential limitations in objectivity inherent in these results, given that the researcher both collected and analysed the data. Additionally, a discussion of the participant count and amount of data is warranted. Future studies may provide valuable insights for more robust conclusions.

This paper shows that online role-play as an approximation of mathematics teaching offers MTE opportunities for pre-service mathematics teachers to practise using different types of formative feedback and to prepare for responding to pupils in the moment. It also underscores the importance of considering specific aspects such as a broader context description for the role-plays and providing the online environment with opportunities to use several forms of expression to optimise the effectiveness of online role-playing. The insights gained from this initial study will inform the design and considerations for a follow-up study. Data were collected outside of the MTE context, highlighting contextual limitations but also prompting considerations for further studies.

Building on the results of this paper and the follow-up study, it may be concluded that role-playing in online MTE could be a topic for future research.

References

Anthony, G. & Walshaw, M. (2009). Characteristics of effective teaching of mathematics: a view from the West. *Journal of Mathematics Education*, 2 (2), 147–164.

Ball, D. L. & Cohen, D. K. (1999). Developing practice, developing practitioners: toward a practice-based theory of professional education. *Teaching as the learning profession: Handbook of policy and practice*, 1, 3–22.

Björklund Boistrup, L. (2022). *Bedömning i matematik pågår: återkoppling för elevers engagemang och lärande* (2 ed.). Liber.

Braun, V. & Clarke, V. (2022). *Thematic analysis: a practical guide*. Sage.

Di Lonardo Burr, S. M., Douglas, H., Vorobeva, M. & Muldner, K. (2020). Refuting misconceptions: computer tutors for fraction arithmetic. *Journal of Numerical Cognition*, 6 (3), 355–377. doi: 10.5964/jnc.v6i3.310

Dyment, J. E. & Downing, J. J. (2020). Online initial teacher education: a systematic review of the literature. *Asia-Pacific Journal of Teacher Education*, 48 (3), 316–333. doi: 10.1080/1359866X.2019.1631254

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E. & Williamson, P. W. (2009). Teaching practice: a cross-professional perspective. *Teachers College Record*, 111 (9), 2055–2100. doi: 10.1177/016146810911100905

Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77 (1), 81–112.

Hayes, S., Capel, S., Katene, W. & Cook, P. (2008). An examination of knowledge prioritisation in secondary physical education teacher education courses. *Teaching and teacher education*, 24 (2), 330–342. doi: 10.1016/j.tate.2006.10.012

Hirsh, Å. & Lindberg, V. (2015). *Formativ bedömning på 2000-talet – en översikt av svensk och internationell forskning*. Vetenskapsrådet.

Howell, H. & Mikeska, J. N. (2021). Approximations of practice as a framework for understanding authenticity in simulations of teaching. *Journal of Research on Technology in Education*, 53 (1), 8–20. doi: 10.1080/15391523.2020.1809033

Hufferd-Ackles, K., Fuson, K. C. & Sherin, M. G. (2004). Describing levels and components of a math-talk learning community. *Journal for Research in Mathematics education*, 35 (2), 81–116. doi: 10.2307/30034933

Kazemi, E., Ghousseini, H., Cunard, A. & Turrou, A. C. (2015). Getting inside rehearsals: insights from teacher educators to support work on complex practice. *Journal of Teacher Education*, 67 (1), 18–31. doi: 10.1177/0022487115615191

Lajoie, C. (2018). Learning to act in-the-moment: prospective elementary teachers' role-playing on numbers. In G. J. Stylianides & K. Hino (Eds.), *Research advances in the mathematical education of pre-service elementary teachers: an international perspective* (pp. 231–243). Springer. doi: 10.1007/978-3-319-68342-3_16

Lampert, M., Franke, M. L., Kazemi, E., Ghousseini, H., Turrou, A. C. et al. (2013). Keeping it complex: using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education*, 64 (3), 226–243. doi: 10.1177/0022487112473837

Matsumoto-Royo, K. & Ramírez-Montoya, M. S. (2021). Core practices in practice-based teacher education: a systematic literature review of its teaching and assessment process. *Studies in Educational Evaluation*, 70. doi: 10.1016/j.stueduc.2021.101047

McDonald, M., Kazemi, E. & Kavanagh, S. S. (2013). Core practices and pedagogies of teacher education: a call for a common language and collective activity. *Journal of Teacher Education*, 64 (5), 378–386. doi: 10.1177/0022487113493807

Radmehr, F. & Goodchild, S. (2022). Switching to fully online teaching and learning of mathematics: the case of Norwegian mathematics lecturers and university students during the covid-19 pandemic. *International Journal of Research in Undergraduate Mathematics Education*, 8 (3), 581–611. doi: 10.1007/s40753-021-00162-9

Rowe, M. B. (2003). Wait-time and rewards as instructional variables, their influence on language, logic, and fate control: part one-wait-time. *Journal of research in science teaching*, 40 (2), 19–32.

Ruiz-Primo, M. & Li, M. (2013). Examining formative feedback in the classroom context. New research perspectives. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 215–232). SAGE. doi: 10.4135/9781452218649

Ryve, A., Nilsson, P., Palm, T., Van Steenbrugge, H., Andersson, C., et al. (2015). *Kartläggning av forskning om formativ bedömning, klassrumsundervisning och läromedel i matematik: delrapport från skolforsk-projektet*. Vetenskapsrådet.

Shaughnessy, M. & Boerst, T. (2018). Designing simulations to learn about pre-service teachers' capabilities with eliciting and interpreting student thinking. In G. J. Stylianides & K. Hino (Eds.), *Research advances in the mathematical education of pre-service elementary teachers: an international perspective* (pp. 125–140). Springer. doi: 10.1007/978-3-319-68342-3_9

Stovner, R. B. & Klette, K. (2022). Teacher feedback on procedural skills, conceptual understanding, and mathematical practices: a video study in lower secondary mathematics classrooms. *Teaching and teacher education*, 110, 103593. doi: 10.1016/j.tate.2021.103593

Van Leeuwen, T. (2005). *Introducing social semiotics*. Routledge.

Crumbs of knowledge – assessing preservice teachers' written probability reasoning

Anna Teledahl, Andreas Bergwall
and Andreas Eckert

Assessment is a complex endeavour, especially in relation to the multi-faceted knowledge required of teachers. In this paper, we investigate two approaches to assessment, analytic and holistic, in the case of pre-service teachers' probability reasoning. Sixteen written student solutions to two exam problems involving conditional probability are analysed. The results show that the analytical approach tends to give a more favourable profile of student knowledge, while essential shortcomings and misconceptions become more explicit with the holistic approach. Results also indicate that how "crumbs of knowledge" are weighed highly influences student knowledge profiles for both approaches. A discussion on how pre-service teachers' mathematical knowledge is assessed should take different assessment approaches into consideration.

Assessment in education has always been a complex endeavour. Assessing, interpreting, grading, or evaluating students' work can be done using different approaches, but regardless of which approach to assessment you use, there are always concerns regarding issues of validity and reliability (Sadler, 2009; Wiliam, 2010). Analytic approaches to assessment infer the identification of specific aspects of the knowledge we want to see, while holistic approaches, albeit noting specific features, view students' work as a whole and respond to it with a global judgment (Sadler, 2009). Analytic approaches have been favoured for identifying and providing teachers and students with detailed information useful for formative feedback (Wisniewski et al., 2020). Holistic approaches are put forward to ensure diversity and relation to the authenticity of tasks while retaining reliability (Walton & Martin, 2023). In assessment in teacher education, there is a need for detailed information to be used for formative purposes but also to guarantee a satisfactory level of achievement. There is also an interest in more authentic assessments strongly related to teachers' practice. Analytic and holistic approaches to assessment are common in teacher education, as

Anna Teledahl, Örebro University
Andreas Bergwall, Örebro University
Andreas Eckert, Örebro University

they are employed in different courses for different purposes. When assessing pre-service teacher's content knowledge, such as mathematics, the practice is similar to the traditional assessment approach in school mathematics (Suurtamm et al., 2016). This practice is analytical and presupposes a decomposition of the subject matter and operationalisation of different levels of mastery concerning the various elements identified in the decomposition.

The point of departure for this study is the assumption that different assessment approaches create grading dilemmas (Sadler, 2013). It can be argued that these dilemmas are particularly pressing in teacher education, not only because the types of knowledge required from pre-service teachers differ from that of other mathematics students but also because the pre-service teachers will pick up elements from the assessment practice that they are subjected to regardless of what they are taught in other situations. The dilemmas present in teacher education concern issues such as validity and reliability, as has been shown by Fauskanger (2015), who identified an inconsistency between the knowledge pre-service mathematics teachers display in two different scoring models. The working hypothesis in this study is that analytic approaches to grading students' mathematics exams will yield different student knowledge profiles compared to holistic approaches. We define a knowledge profile as a comprehensive assessment of a student's mathematics knowledge, skills, and abilities (Segers et al., 1999). However, we have yet to identify how they potentially differ and whether the difference matters. This study explores these questions by employing and comparing an analytic and holistic approach to assessing preservice teachers' mathematics exam tasks. We aim to add to the body of knowledge on assessment in mathematics, specifically in mathematics teacher education. We use the mathematical topic of conditional probability as our case; the research question for the paper thus is: *What is the difference between the knowledge profiles created by using analytic and holistic approaches in assessing preservice teachers' knowledge concerning conditional probability?*

Background

Assessment is as old as education and has several purposes (Newton, 2007). When used to determine students' level of achievement or competence in relation to norms, criteria, or standards, it is often termed summative. Formative assessment provides students and teachers with detailed information on their learning progression and how and what they can improve (Wisniewski et al., 2020).

The history of assessing students' achievement is long and includes large-scale assessment, which traditionally presumes a psychometric perspective, as well as classroom assessment, which views assessment as a social practice. (Suurtamm et al., 2016). The psychometric perspective in large-scale assessment has traditionally dealt with assessment as a question of measurement's reliability, which can lead to formats that favour single right-answer questions. Such testing practices reduce school subjects like mathematics to manageable, unambiguous content areas. It has been pointed out that such a perspective fits poorly with modern ideas of mathematics learning (Suurtamm et al., 2016; Wiliam, 2010). Other dilemmas with assessment in mathematics concern the way assessment is graded. Holistic approaches look at students' work as a whole and arrive directly at a judgment. In contrast, analytic approaches make separate judgments on the different parts of a student's work, measured against criteria (Sadler, 2009). Analytic assessment relies on rating scales or rubrics that operationalise criteria and provide arithmetic models for calculating a grade. When different aspects of knowledge are disassembled and articulated as criteria in the form of rubrics, schemes, grids, or matrices, this type of assessment is often labelled criterion-based (Sadler, 2009). Holistic scoring schemes ”associate each grade level with a reasonably full verbal description, which is intended as indicative rather than definitive or prescriptive” (Sadler, 2009, p. 6). Sadler argues that analytic grading approaches are the most common in recent years, especially in higher education, but he suggests that they give rise to what he calls ”anomalies.” An example of an anomaly is when teachers discover that their intuitive impressions of student work are at odds with the outcome of an analytic scoring approach. They can find that work they consider brilliant does not meet the criteria when you look at its parts or that work teachers find mediocre can come off as very good according to the separate analytic scoring. A second anomaly is when teachers perceive that the criteria are incomplete and find themselves wishing for an extra criterion that would capture the quality, they believe they can identify or think is missing in a student's work.

Assessing preservice teachers requires a grading teacher to attend to validity issues because the construct of mathematics knowledge is defined differently than for other mathematics students (Fauskanger, 2015). Preservice teachers are expected to attain a combination of different types of knowledge, of which subject matter knowledge is only one (Loewenberg Ball et al., 2008). Even if the assessment should focus exclusively on the subject matter, there are different types of subject matter knowledge, for example, knowledge of facts and procedures and knowledge of concepts and connections (Tchoshanov, 2011), which puts a particular demand on the assessment. Another issue is that knowledge constructs vary between mathematics topics (Copur-Gencturk et al., 2022; Wiliam, 2010). Probability is an example of a mathematical topic that differs from others in that it is not purely deterministic. Chance makes people intuitively lean towards subjective reasoning that does not necessarily align with mathematics (Kahneman et al., 1982). Studies show, for example, how students disregard sample size (Kahneman et al., 1982), underlying probability distributions (Kahneman et al., 1982; Lecoutre, 1992), and dependent/independent events (Kahneman et al., 1982). These tendencies are not limited to young learners. Studies also show that in-service and preservice teachers often adhere to the same faulty subjective probability reasoning as students (e.g. Batanero & Diaz, 2012).

For the case of assessing preservice teachers' probabilistic reasoning, we build on Jones et al. (1999), in which probabilistic thinking refers to "thinking in response to any probability situation" (p. 488). A probability situation is any activity or experiment where the outcome cannot be predetermined precisely. Further, probabilistic thinking is characterised in relation to four content areas: *sample space, probability of an event, probability comparisons* and *conditional probability*. Students' probabilistic thinking is assumed to develop over time in all four areas. Four levels of development are identified: *subjective* (Level 1), *transitional* (Level 2), *informal quantitative* (Level 3), and *numerical* (Level 4). Jones et al. (1999) characterise them as what can be observed when students reason about probability situations. For that reason, we refer to them as levels of *probabilistic reasoning*. The details are presented in table 1. As our data do not include situations where students compare different sample spaces, we have omitted the part of the framework that relates to probability comparisons.

Table 1. *Framework for probabilistic reasoning based on Jones et al. (1999)*

	Level 1	Level 2	Level 3	Level 4
Sample space	Lists an incomplete set of outcomes for a one-stage experiment.	Lists a complete set of outcomes for a one-stage experiment. Sometimes lists a complete set of outcomes for a two-stage experiment using limited and unsystematic strategies.	Consistently lists the outcomes of a two-stage experiment using a partially generative strategy.	Adopts and applies a generative strategy that enables a complete listing of the outcomes for two- and three-stage cases.
Probability of an event	Predicts most/least likely events based on subjective judgments. Recognizes certain and impossible events.	Predicts most/least likely events based on quantitative judgments but may revert to subjective judgments.	Predicts most/least likely events based on quantitative judgements, including situations involving non-contiguous outcomes. Uses numbers informally to compare probabilities. Distinguishes certain, impossible, and possible events and justifies choice quantitatively	Predicts most/least likely events for single-stage experiments. Assigns a numerical probability to an event (either an actual probability or a form of odds).
Conditional probability	Following one trial of a one-stage experiment does not give a complete list of outcomes, even though a complete list was given before the first trial. Recognizes when certain and impossible events arise in non-replacement situations.	Recognizes that the probabilities of some events change in a non-replacement situation; however, recognition is incomplete and is usually restricted to previously occurring events.	Can determine changing probability measures in a non-replacement situation. Recognizes that the probabilities of all events change in a non-replacement situation.	Assigns numerical probabilities in replacement and non-replacement situations. Distinguishes dependent and independent events.

The framework was initially designed to assess young learners. However, we use it in this paper as a baseline or minimum requirement to assess preservice subject matter knowledge.

Method

Data sample

Our data consists of preservice teachers' solutions to two tasks from written exams on a mathematics education course for pre-service middle school teachers.

> Task I: On Christmas Eve, Santa Claus arrives with three gifts each for Otto and Vira in his sack. He puts the sack down, bends down, and randomly picks up two gifts. What is the probability that one is for Otto and one is for Vira?
>
> Task II: Otto and Vira are going to buy lottery tickets. There are 100 tickets, but only 10 of them are winning tickets. "The probability of getting a winning ticket is only 10 %," says Vira. "Yes, but if we buy two, it becomes twice as big, 20 %," says Otto. "And if we buy ten tickets, it becomes ten times as big," says Vira. "Then we know we'll win!" Are Otto and Vira right? Explain your thinking carefully!

The tasks are mathematically similar in that they concern two-stage non-replacement situations. They can be solved using a standard strategy: A complete list of outcomes can be generated and represented using a tree diagram. Probabilities of the relevant two-stage outcomes can then be determined by multiplying the probabilities along the corresponding branches of the tree and then adding them to give the probabilities of the events described in the tasks.

The tasks also have significant differences. Task I asks for a specific numerical probability. Due to its symmetry, it has a simple solution that does not require a tree diagram: Regardless of whether Otto or Vira gets the first gift, three of the remaining five are for the other one, and the answer must be 3/5. There is no such symmetry in task II. However, task II does not ask for a probability but for evaluating statements, some representing common misconceptions. It is worth noting that had there been only ten lottery tickets and only one winning ticket, Otto's and Vira's reasoning in task II would have been correct.

Task I was part of a regular exam during a mathematics course in teacher education, in which 46 students participated. Task II was part of the corresponding re-exam, i.e. only students who failed or did not participate in the first exam could attend. 24 students attended the second exam. Both exams were anonymous. Solutions produced by all students who participated in both exams were sought out and paired together by a study administrator. In total, this yielded 16 student solutions (for both tasks).

Method of analysis

Individually, we looked through all 16 solutions and chose five that represented a wide range of ways to approach the tasks, solve them, and communicate the solutions, as well as varying levels of probability reasoning. We compared our choices, agreed upon five student solutions for a deeper analysis, and produced short descriptions of each solution, including similarities and differences between task I and II. We assessed the students' levels of probability reasoning using analytic and holistic approaches. In the analytic analysis, we looked for individual smaller units of knowledge from the Framework for probabilistic reasoning based on Jones et al. (1999), where connections between the units were not considered. An example of such a unit is ”The student can determine changing probability measures in a non-replacement situation.”. The holistic analysis sought to find connections between various signs of knowledge to create a more comprehensive description of what the student ”knows”. An example of this is ”the student shows a good understanding of sample spaces in one and two-stage events” where there is a clear connection between one and two-stage events. The findings were summarised in ”knowledge profiles” for each student. From the analytic perspective, we sought evidence of the highest levels of probabilistic reasoning exhibited. For instance, if (part of) a student's solution indicated probabilistic reasoning on level 3 in a particular content area, the student's reasoning was considered to be on that level even if other parts of the solutions only indicated lower levels of reasoning. For the holistic perspective, we searched for evidence indicating that a certain level had been reached or not reached. We combined them to create holistic knowledge profiles of the students' probabilistic reasoning.

Results

Our analysis shows that adhering to different assessment perspectives on students' written answers results in varying knowledge profiles. Our analytical approach proved to produce more favourable profiles than our holistic one. Our data does not enable us to evaluate which is more accurate. However, it discusses how different perspectives affect our understanding of students' mathematical knowledge in teacher training.

Figure 1 shows a reproduction of two students' written solutions, chosen to illustrate similarities and differences between an analytic and a holistic assessment. On the left, Student A produced a tree diagram with clear connections to the gift context in task I with colour coding and a bag drawing. All the facts, conditions, and assigned probabilities included in the answers are correct. The reasoning in task II, lower left corner, stops short of what is asked for in the task, but everything written is correct. Student B, on the right, also uses a tree diagram in task I but lacks explanations, and the assigned probabilities are incorrect for some events in the lower branches. Student B's reasoning in task II, lower right corner, starts correctly but soon focuses on the event of drawing one particular lottery ticket rather than a winning ticket. Student B ends the reasoning with an analogy to task II, with conditions similar to task I, but suggests a replacement situation rather than the non-replacement situation that task II is originally about.

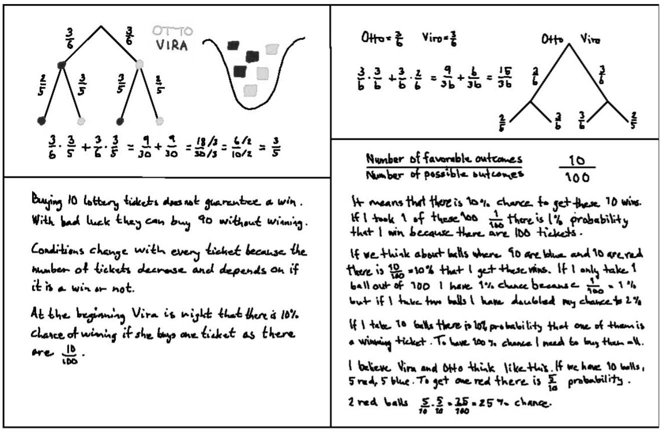

Figure 1. *Solutions to task I (top) and II (bottom) produced by Student A (left) and B (right), translated and reproduced by the authors*

We summarise our assessment of Student A's written answers in table 2. The student scores top marks within each category in the analytical assessment. The knowledge profile does not reflect that the student does not solve task II completely. We conclude that this student can reason appropriately on a conditional probability task as assessed from an analytical perspective based on the chosen framework. The student receives a positive assessment from a holistic assessment as well. However, the knowledge profile now includes the absence of a complete answer in task II. Less tangible clues, such as clear drawings and the absence of information, elevate and demote the holistic assessment. Since the answer from task II lacks the elegant tree diagram representation and calculation of the conditional probabilities, the holistic assessment indicates that the student's reasoning ability is limited to familiar tasks. If task II represents an unfamiliar probability situation for the student, they refrain from speculating or stepping into unknown territory. Everything written is correct, possibly indicating a recognition of the unfamiliarity and an inclination to stick to what is known. As with most holistic assessments, ours requires interpretations of what the student knows and is able to do, and as with every assessment, there is no telling if the student does not know or fails to understand something just because their written answers do not show this.

Table 2. *Assessment of Student A's written answers for task I and II*

		Analytical ass.	Holistic assessment
Sample space	Reasoning that includes a complete sample space of a two-stage event but only in task I. The lottery context of task II did not elicit the same reasoning. However, it stopped at recognising that it was about dependent events and presented the sample space of the one-stage event. The written answer includes a clear connection between the representation of the sample space, the tree diagram in task I and a list in task II, and the context.	4 – Can present a tree diagram with all possible outcomes for task I.	The written answer indicates a good understanding of sample spaces in one and two-stage events.
Prob. of an event	Reasoning includes assigning numerical probabilities for one-stage and two-stage events. The reasoning distinguishes between certain and possible events.	4 – Can correctly assign numerical probabilities for one-stage events in task I and II.	The written answer indicates a good understanding of task I, which is numerical probabilities for one- and two-stage events. However, by stopping at distinguishing between different events in task II instead of calculating numerical probabilities, the student fails to display a comprehensive understanding of the difference between the two tasks.
Cond. prob.	Assigns numerical probabilities in non-replacement situations in task I but not task II. The student stops at recognising that the probabilities of all events change in a non-replacement situation in task II.	4 – Can assign numerical probabilities in non-replacement situations in task I.	The difference in handling non-replacement situations in task I and II indicates that students can only reason successfully when the context and conditions are familiar. However, everything stated is correct and no misunderstandings are put forward.

We summarise our assessment of Student B's written answers in table 3. The student scores top marks within the Probability of an event but only a 2 in Sample space and Conditional probability. Student B succeeds with assigning probabilities for the first stage in both tasks I and II, as seen in the very beginning of both written answers. It is possible to look at the number of branches in the tree diagram and interpret it as representing the complete sample space. However, the lack of labels and clarity in the presentation and the incorrect probabilities mean that such an interpretation goes far beyond what is shown in the answer. This leads to a low mark in the Sample space. The fact that half of the probabilities for stage two are incorrect results in a low mark in Conditional probability. The holistic assessment of sample space in tasks I and II is that the student has strategies for listing complete sample spaces and can apply them in some context. Multiple instances lead to this conclusion since the student must understand the complete sample space for the presented reasoning. The main issue lies with the conditional probability reasoning; it is unclear whether Student B recognises the difference between replacement and non-replacement situations and how all probabilities are affected in non-replacement tasks.

In the case of Student A, the analytic assessment led to a favourable knowledge profile within all areas compared to the holistic one. The holistic

Table 3. *Assessment of Student B's written answers for tasks I and II*

		Analytical assessment	Holistic assessment
Sample space	Reasoning that includes a sample space of a two-stage event but only in task I. The student uses a tree diagram with incoherent labels and fails to use the correct branches in the following calculations. The chosen representation in the written answer has minimal explicit connections to the context of task I. The lottery context of task II did not elicit the tree-diagram strategy.	2 – Can list a complete set of outcomes for a one-stage experiment and use a limited and unsystematic strategy (a faulty tree diagram) to list a set of outcomes for a two-stage experiment in task I.	The written answer indicates an okay understanding of sample spaces in one and two-stage events, at least in some contexts. However, the presentation lacks precision and clarity.
Prob. of an event	Reasoning that includes assigning numerical probabilities for one-stage events in both tasks. The reasoning distinguishes between certain and possible events. The reasoning in task II includes several instances of assigning numerical probabilities for one-stage events but not always the ones connected to the task.	4 – Can correctly assign numerical probabilities for one-stage events in task I and II.	The written answer indicates an understanding of numerical probabilities for one-stage events in task I and II. However, the expanded reasoning in task II indicates that the student is struggling to differentiate between drawing a specific ticket and a winning ticket.
Conditional prob.	Assigns numerical probabilities correctly in replacement situations in task II even though that was not the task. The reasoning includes that probabilities of some events change in a non-replacement situation but fails to recognise how all events' probabilities change in task I and II.	2 – Can recognise that the probabilities of some events change in a non-replacement situation in task I; however, recognition is incomplete and is restricted to events that have previously occurred	The example used in the reasoning for task II is a replacement situation, and the task is about a non-replacement situation. However, correctly handling the replacement situation indicates that the student exhibits at least some understanding of conditional probability.

assessment, however, offers a more nuanced profile by considering what is not written. There are several indications of confidence in many areas and an unwillingness to venture into unknown territory with speculations. The inability to recognise task II as a non-replacement situation, the same as in task I, is only picked up in the holistic approach. Regarding Student B, the analytic assessment led to a less favourable knowledge profile in two out of three areas than Student A. The holistic assessment nuances the profile by combining several crumbs of knowledge that are identified even when the calculations are incorrect. This may lead to a split conclusion that can be interpreted as both too positive (there are only crumbs of knowledge found in incorrect calculations) and too negative (there is, in fact, evidence of some understanding), depending on your perspective. The two contrasting approaches yield different knowledge profiles, but these also depend on the weight given to the crumbs of knowledge.

Discussion

The difference between the knowledge profiles confirms that the two assessment approaches do not align. It may be worth pointing out that although this paper shows only two of the analysis's knowledge profiles, the conclusions align with the rest. We can argue that two of Sadler's (2009) anomalies are relevant to our cases. We see some crumbs of knowledge displayed by Student B that are not picked up by the analytic assessment. In contrast, we can argue that the analytic assessment of Student A produces a knowledge profile that may be too positive since the student fails to recognise the mathematical similarities between task I and II.

The ability to disregard context, recognise mathematical similarities between situations, and recognise how their interpretations align with is essential for teachers (Loewenberg Ball et al., 2008). It is important to note that neither of our two assessment approaches attended to what the students failed to do, although this was at least acknowledged in the holistic approach. This represents a typical dilemma in assessment where we can only infer understanding from what is done, not misunderstanding or lack of understanding from what is not done. As mentioned above this is especially salient in the case of student A. However, because this student does not speculate but stops while they are ahead, their profile is viewed more favourably in the holistic approach.

The practice of recognising what we have referred to as crumbs of knowledge, i.e. inferring knowledge from one correct example while ignoring several examples of lack of success with tasks concerning the same mathematical idea, is common. This represents a challenge in all assessments but becomes critical in teacher education, where it can be argued that the students must display a thorough understanding of topics such as probability (Tchoshanov, 2011). It is challenging for teacher educators to weigh the importance of different crumbs of knowledge found in the solutions to produce a highly valid and reliable student knowledge profile. Failing to detect misconceptions such as the one displayed by Student B is a high risk with analytic approaches as well as holistic approaches, where crumbs of knowledge are allowed to compensate for other shortcomings. Assessment based on crumbs and ”what is done” seems to increase the risk of missing common preservice teachers' misconceptions identified by, e.g. Batanero and Diaz (2012). It is possible that a combination of analytic and holistic approaches can compensate for this risk. However, discussing how to weigh crumbs of knowledge may also be essential in securing high reliability and validity. Considering this limited example, we call for further studies to provide more details about assessment anomalies in teacher education and how they might affect preservice teachers' future (assessment) practices.

References

Batanero, C. & Diaz, C. (2012). Training school teachers to teach probability: reflections and challenges. *Chilean Journal of Statistics*, 3 (1), 3–13.

Copur-Gencturk, Y., Jacobson, E. & Rasiej, R. (2022). On the alignment of teachers' mathematical content knowledge assessments with the common core state standards. *Journal of Mathematics Teacher Education*, 25 (3), 267–291. doi: 10.1007/s10857-021-09486-4

Fauskanger, J. (2015). Challenges in measuring teachers' knowledge. *Educational Studies in Mathematics*, 90 (1), 57-73. doi: 10.1007/s10649-015-9612-4

Jones, G. A., Langrall, C. W., Thornton, C. A. & Mogill, A. T. (1999). Students' probabilistic thinking in instruction. *Journal for Research in Mathematics Education*, 30 (5), 487–519. doi: 10.2307/749771

Kahneman, D., Slovic, P. & Tversky, A. (1982). *Judgement under uncertainty: heuristics and biases*. Cambridge University Press.

Lecoutre, M.-P. (1992). Cognitive models and problem spaces in ”purely random” situations. *Educational Studies in Mathematics*, 23 (6), 557–568. doi: 10.1007/BF00540060

Loewenberg Ball, D., Thames, M. H. & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59 (5), 389–407. doi: 10.1177/0022487108324554

Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education: Principles, Policy & Practice*, 14 (2), 149–170. doi: 10.1080/09695940701478321

Sadler, D. R. (2009). Transforming holistic assessment and grading into a vehicle for complex learning. In G. Joughin (Ed.), *Assessment, learning and judgement in higher education* (pp. 1–19). Springer. doi: 10.1007/978-1-4020-8905-3_4

Sadler, D. R. (2013). Assuring academic achievement standards: from moderation to calibration. *Assessment in Education: Principles, Policy & Practice*, 20 (1), 5–19. doi: 10.1080/0969594X.2012.714742

Segers, M., Dochy, F. & De Corte, E. (1999). Assessment practices and students knowledge profiles in a problem-based curriculum. *Learning Environments Research*, 2 (2), 191–213. doi: 10.1023/A:1009932125947

Suurtamm, C., Thompson, D. R., Kim, R. Y., Diaz Moreno, L., Sayac, N. et al. (2016). State of the art. In C. Suurtamm, D. R. Thompson, R. Y. Kim, L. Diaz Moreno, N. Sayac et al. (Eds.), *Assessment in mathematics education large-scale assessment and classroom assessment* (pp. 2-29). Springer.

Tchoshanov, M. A. (2011). Relationship between teacher knowledge of concepts and connections, teaching practice, and student achievement in middle grades mathematics. *Educational Studies in Mathematics*, 76 (2), 141–164. doi: 10.1007/s10649-010-9269-y

Walton, J. & Martin, J. L. (2023). Applying Sadler's principles in holistic assessment design: a retrospective account. *Teaching in Higher Education*, 1–18. doi: 10.1080/13562517.2023.2244439

Wiliam, D. (2010). What Counts as evidence of educational achievement? The role of constructs in the pursuit of equity in assessment. *Review of Research in Education*, 34 (1), 254–284. doi: 10.3102/0091732x09351544

Wisniewski, B., Zierer, K. & Hattie, J. (2020). The power of feedback revisited: a meta-analysis of educational feedback research. *Frontiers in Psychology*, 10. doi: 10.3389/fpsyg.2019.03087

Collaboration between mathematicians and mathematics educators in secondary teacher education

Olov Viirman, Yukiko Asami-Johansson,
Jonas Bergman Ärlebäck, Mikael Cronhjort
and Magnus Jacobsson

In this symposium, three research and development projects from different Swedish universities will serve as a starting point for a discussion of different ways in which mathematicians and mathematics educators can work together to improve the preparation of prospective secondary mathematics teachers for teaching mathematics.

Comparatively little research has been focused on the education of prospective secondary, and in particular upper secondary, mathematics teachers (USMT's). There is some consensus among mathematicians and teacher educators (e.g., Dreher et al., 2018; Leikin et al., 2018) that prospective USMTs need solid knowledge of central topics in secondary mathematics, such as calculus and algebra, including some more advanced topics. However, research on USMT education often takes the content and format of mathematics subject matter courses as given, and separates them from the mathematics education courses that are the main focus of the research.

However, many prospective and practicing mathematics teachers fail to see the relevance of their higher-level mathematics studies to their teaching practice (e.g., Wasserman et al, 2018; Zazkis & Leikin, 2010). It has been argued (e.g., Wasserman et al, 2018; Winsløw & Grønbæk, 2014) that one possible explanation is that the higher-level mathematics courses taken by prospective teachers rarely discuss the content from a didactical perspective. In the light of this, research has begun investigating the role that mathematics courses, and the mathematicians who teach them, play in USMT education. For instance, Yan et al. (2020) address the question of how to better connect the mathematics studies of prospective USMTs to their future teaching practice, and Alvarez et al. (2020) present principles for designing tasks that address applications in secondary mathematics education, for use in undergraduate mathematics

Olov Viirman, Uppsala University
Yukiko Asami-Johansson, University of Gävle
Jonas Bergman Ärlebäck, Linköping University
Mikael Cronhjort, University of Gävle
Magnus Jacobsson, Uppsala University

courses. A recent special issue of ZDM (Buchbinder et al., 2023) is devoted to ”exploring and strengthening university mathematics courses for secondary teacher preparation”, where many of the studies touch on the idea of collaboration between mathematicians and didacticians, either in practice or in research. In the Swedish context, however, there is little research on USMT education in general, and even less on what form such collaboration might take. In what follows, we will present three projects from different Swedish universities that address this question in different ways.

Project 1 – An advanced course in mathematics and didactics for USMTs

One way to understand the demands of USMT education is through the so-called ”double discontinuity”, first formulated by the mathematician Felix Klein (e.g., Winsløw & Grønbæk, 2014). The first discontinuity occurs when students enrol at university to become mathematics teachers. The mathematics they then encounter is very different from what they have experienced in school, the difference becoming more pronounced as their studies progress. When they return to schools as teachers, the second discontinuity occurs, as they are expected to teach the ”school mathematics” they left behind when entering university. With this as a starting point, and drawing on related initiatives elsewhere (Wasserman et al., 2017; Winsløw & Grønbæk, 2014), at Uppsala University we have developed an interdisciplinary course in advanced mathematics and didactics (Viirman & Jacobsson, 2022), that covers parts of real analysis and abstract algebra relevant to upper secondary mathematics and treats mathematics and didactics in parallel and together. It is taught jointly by a mathematician and a didactician through so-called ”team teaching”, with both teachers present in the classroom during all teaching sessions. The course has run twice so far, and in the symposium we will present examples of how the collaboration has created opportunities for learning. For instance, the interdisciplinarity enables us to address not only local relevance, where the same content is taught at university and in school, but also non-local relevance, where specific mathematics content can serve as examples of general didactic relevance, for example concerning the role of examples or classification. Co-teaching allows, among other things, for didactic reflection on the teaching as it happens.

Project 2 – Mathematical concept formation for USMTs

At the University of Gävle, we have revised the teaching of the course ”Mathematical concept formation for subject teachers”. The course has been given by three teachers in cooperation: a didactician, a mathematician, and a theoretical physicist who has focused on didactics during the last decades. Teachers change their institutional positions (regardless of their ”actual” positions) in in relation to the emergence of knowledge (Chevallard, 2021). That is, a mathematician may change position to teacher educator, or a didactician to teacher educator, depending on what the didactic stakes of the lesson are. Such an awareness of our own institutional positions enabled fruitful discussions and new perspectives during our revision of the course. When transposing the knowledge for teaching mathematics to USMTs, there are two dimensions, namely, the learning of mathematics as a domain (didactic level), and learning the epistemology of didactics (paradidactic level) (Otaki & Asami-Johansson, 2022). On a didactic level, mathematicians perceive the position of prospective teachers as students and focus on their need to master, for example, mathematical concepts. On a paradidactic level, teacher educators (as an institutional position) perceive the position of prospective teachers as teachers, focusing on their ability to teach future students the mathematical concepts. With these two dimensions in mind, especially the paradidactic level, a guiding principle in our revision was to include mathematical concepts that are known from research and experience to be frequently misunderstood. These include the number system, logical reasoning, proofs, powers, roots, logarithms, and digital tools, including GeoGebra. We also included the concept of distance with some different metrics, as this would allow for examples that we expected would be unfamiliar to the prospective teachers. For example, we used the Manhattan metrics to demonstrate that a circle does not look like the Euclidian circles we are used to. This was based on materials developed at Kleindagarna, which is an arena for meetings between mathematicians and upper secondary mathematics teachers, inspired by Felix Klein (Cronhjort & Hagemann-Jensen, 2021). In the symposium, we will present examples from Kleindagarna of how higher mathematics can enrich conceptual understanding in school mathematics.

Project 3 – Theoretical foundations of secondary school calculus

At Linköping University, the undergraduate courses in mathematics (large lecture settings) and mathematics education (more intimate class settings) are generally taught separately but in parallel. However, two courses integrate mathematics and didactics: a course on mathematical modelling and ”Calculus: Theoretical Foundations for Secondary Mathematics Teaching, 4hp”. The main aims of the latter, newly developed course are (i) to provide a theoretical foundation for parts of secondary school mathematics and (ii) to give students the opportunity to relate and practice the content in a way that is relevant to their own future teaching practice. The course actively seeks to engage students in understanding and exploring their dual roles as learners and teachers of mathematics (cf. Bowers & Doerr, 2001). The mathematical content focuses on fundamental theorems about sequences, supremum (infimum), limits, functions, (uniform) continuity, derivatives and integrals. The course examination is closely linked to ways of working as a practising teacher: (i) presenting solutions orally and in writing; (ii) providing comments and feedback on (fellow students') solutions; and (iii) designing a GeoGebra activity intended for the upper secondary level that illustrates a central idea/notion/concept/technique of the course. In the case of this course, the developer and the teacher of the course are one and the same person, and his dual role as a mathematician and a didactician will be discussed and problematised in this symposium.

Concluding remarks

The projects presented above have obvious similarities but also display clear differences. They all seek to engage prospective teachers in reflection on their dual roles, both as students of mathematics and as future teachers. However, the institutional roles of the academics involved in the projects are different, which has implications for how the collaboration takes shape in the different settings. In the symposium, we aim to use these projects as starting points for a deeper discussion about how to improve USMT education in Sweden, and about the role that mathematicians and didacticians can play in this work. We believe that such a discussion is long overdue within the Swedish mathematics education community.

References

Álvarez, J. A., Arnold, E. G., Burroughs, E. A., Fulton, E. W. & Kercher, A. (2020). The design of tasks that address applications to teaching secondary mathematics for use in undergraduate mathematics courses. *The Journal of Mathematical Behavior*, 60, 100814.

Bowers, J. & Doerr, H. M. (2001). An analysis of prospective teachers' dual roles in understanding the mathematics of change: eliciting growth with technology. *Journal of Mathematics Teacher Education*, 4 (3)1, 15–137.

Buchbinder, O., Wasserman, N.H. & Buchholtz, N. (2023). Exploring and strengthening university mathematics courses for secondary teacher preparation [Special issue]. *ZDM*, 55 (4).

Chevallard, Y. (2022). Challenges and advances in teacher education within the ATD. In Y. Chevallard, B. Barquero, M. Bosch, I. Florensa, J. Gascón, P. Nicolás & N. Ruis-Munzon. (Eds.), *Advances in the anthropological theory of the didactic* (pp. 81–89). Birkhäuser.

Cronhjort, M. & Hagemann-Jensen, M. (2021). Avstånd med olika metriker. *Nämnaren*, 2021 (1), 44–50.

Dreher, A., Lindmeier, A., Heinze, A. & Niemand, C. (2018). What kind of content knowledge do secondary mathematics teachers need? *Journal für Mathematik-Didaktik*, 39 (2), 319–341.

Leikin, R., Zazkis, R. & Meller, M. (2018). Research mathematicians as teacher educators: focusing on mathematics for secondary mathematics teachers. *Journal of Mathematics Teacher Education*, 21 (5), 451-473.

Otaki, K. & Asami-Johansson, Y. (2022). Exploring the paradidactic ecosystem: conditions and constraints on the teaching profession. In Y. Chevallard, B. Barquero, M. Bosch, I. Florensa, J. Gascón, P. Nicolás & N. Ruis-Munzon (Eds.), *Advances in the anthropological theory of the didactic* (pp. 155-163). Birkhäuser.

Viirman, O. & Jacobsson, M. (2022). Teacher learning from team teaching an advanced course in mathematics and didactics for prospective upper secondary teachers. In M. Trigueros, B. Barquero, R. Hochmuth & J. Peters (Eds.), *Proceedings of INDRUM 2022 – fourth conference of the International Network for Didactic Research in University Mathematics Education* (pp. 548–557). University of Hannover & INDRUM.

Wasserman, N. H., Fukawa-Connelly, T., Villanueva, M., Mejia-Ramos, J. P. & Weber, K. (2017). Making real analysis relevant to secondary teachers: building up from and stepping down to practice. *Primus*, 27 (6), 559–578.

Wasserman, N., Weber, K., Villanueva, M. & Mejia-Ramos, J. P. (2018). Mathematics teachers' views about the limited utility of real analysis: a transport model hypothesis. *The Journal of Mathematical Behavior*, 50, 74–89.

Winsløw, C. & Grønbæk, N. (2014). Klein's double discontinuity revisited: contemporary challenges for universities preparing teachers to teach calculus. *Recherches en Didactique des Mathématiques*, 34 (1), 59–86.

Yan, X., Marmur, O. & Zazkis, R. (2020). Calculus for teachers: perspectives and considerations of mathematicians. *Canadian Journal of Science, Mathematics and Technology Education,* 20, 355–374.

Zazkis, R. & Leikin, R. (2010). Advanced mathematical knowledge in teaching practice: perceptions of secondary mathematics teachers. *Mathematical Thinking and Learning*, 12 (4), 263–281.

Workshops

Teachers' mathematics textbook awareness

Daniel Brehmer

Mälardalens Universitet

Even if the textbook is the most important artefact for planning and teaching mathematics, teachers are rarely taught how to understand and effectively use textbooks for planning and teaching mathematics. Thus, teachers often lack in "awareness" of the potential support mathematics textbooks can provide. In this workshop, we discuss what to include in an education for "mathematics textbook awareness" for teachers (as professional development) and for prospect teachers (in teacher education). We also discuss how to measure if and how teachers' view and use of mathematics textbooks change from such education. The discussions depart from the idea of educative curriculum materials, design principles for such teacher guides and what types of knowledge such material could mediate to teachers and prospect teachers.

Collages as a research methodology in mathematics education research?

Petra S Källberg[1], Ulrika Ryan[1], Anna Chronaki[1], Tamsin Meaney[2] and Eva Norén[3]

[1]Malmö University, [2]Western Norway University of Applied Sciences,

[3]Stockholm University

The aim of this workshop is to explore how collages could function as a research methodology that takes seriously power issues connected to hierarchies, by increasing multivocality and pluralism, in order to facilitate dialogue. In small groups, the participants will explore and respond to some collages that teacher educators have created about their desires for mathematics education in migration contexts. The workshop will focus on the following questions, a) how and what desires can be interpreted in the collages? and b) therefore are collages an ethical research method for co-constructing dialogues?

The mathematics is MInE workshop: developing a teaching model for inclusion and equity

Helena Roos and Anette Bagger

Malmö University and Örebro University

The purpose of this workshop is to present and elaborate on a teaching model designed to facilitate inclusion and equity in mathematics education. Stemming from "the mathematics is MInE project" and influenced by Ainscow's inclusion and equity framework, the model encompasses core elements for mathematics education to benefit every student. By discussing the core elements, addressing common dilemmas, and exercising professional judgment, we aim to develop the model in order to provide a more inclusive and equitable mathematics education.

Short communications

Teaching quality during mathematics lessons

Jöran Petersson[1] and Jorryt van Bommel[2]

[1]Malmö University, [2]Karlstad University

This presentation focuses changes in elements of teaching quality during three parts of a mathematics lesson; start, middle and end. 64 grade 7 lessons were analysed using these partitions, forming a time series that shows decreasing quality for most elements (e.g., Modelling) from start to end of a lesson. Few elements showed an increase of quality (e.g., Intellectual Challenge). The structure of the lessons (whole class at start and individual work during middle and end) may partially account for a possible decrease. Even so, a decrease in quality is not desirable and implications on students' learning are of concern.

Students' perceptions of how feedback affects their mathematical knowledge and test anxiety

Annika Pettersson

Kristinehamn Municipality and Karlstad University

This presentation explores students' perceptions of a teacher-developed feedback method. Many adults in Sweden attend mathematics at Municipal Adult Education to qualify for further studies, but a significant number fail to obtain a degree. Providing specific feedback can help students to succeed. In this action research project, students receive tailored materials that address topics they failed in written tests. After studying the materials, they can try again. They can write or explain their solutions to the tasks orally directly to the teacher. Semi-structured interviews were conducted with ten adult students. Thematic content analysis revealed that students perceived improved performance on previously failed topics and reduced test anxiety.

AI literacy for teaching social sustainability may require specific societal awareness

Christian H. Andersson

Malmö University

This presentation reports on tentative research results indicating that mathematics teachers need specific kinds of societal awareness to facilitate a classroom discussion on ethical and critical perspectives related the mathematics in the 4th industrial revolution (Big Data, AI, Machine learning, etc.). This awareness relates to societal phenomena that may intersect with mathematical modelling, e.g. data may portray whiteness/sexism, risking mathematical analysis to reproduce them. To facilitate classroom discussions on the intersection, teachers must first be aware of how such phenomena operate in society. Teaching units in Sweden and USA are presented with examples of how mathematics teachers enable discussions through their societal awareness. An implication is that teacher education may require more societal content.

Teacher students' identity as mathematics practitioners

Laura Fainsilber

Chalmers university of technology and University of Gothenburg

This presentation investigates how graduating secondary school mathematics teacher students describe the evolution of their relation to mathematics during a 5-year teacher education program. It builds upon in-depth interviews with students. We examine identity in the sense of Sfard and Pruzak, i.e. narratives about individuals that are reifying, endorsable, and significant. These narratives about the students' development express a deepened interest for the subject, a view of the field as a connected whole, an appreciation for a variety of representations, ease with mathematical language and notations, and self-confidence in problem solving. Narratives also show that the process of development builds upon many small steps.

Conceptual knowledge of *addition* and *subtraction* when using length models inspired by the El'konin Davydov curriculum

Lotta Wedman[1] and Helena Eriksson[2]

[1]Umeå University; [2]Dalarna University

It can be hard to teach mathematics in a way that facilitates a development of conceptual knowledge. In this presentation we focus on a pilot study which aimed to explore what conceptual knowledge of addition and subtraction, specifically involving the inverse relationship between them, that can be identified when using length models, inspired by the El'konin Davydov curriculum, in teaching students at grade 4 and 5. Three examples of a length model were given to two teachers, who planned a lesson that was observed and video recorded. The analysis used two dimensions: different qualitative levels of how conceptual knowledge of addition and subtraction is communicated, and implicit vs explicit knowledge. Implicit knowledge of all levels could be found in the communication, but explicit knowledge was not as common.

Students' voices on online enrichment in mathematics

Elisabet Mellroth[1,2] and Jorryt van Bommel[1]

[1]Karlstad University; [2]City of Karlstad

This presentation reports on a study on upper secondary school students' opinions and views regarding an online enrichment in mathematics. Interviews were conducted with seven participating students, who express that the enrichment deepened their mathematical knowledge, in line with the fact that learning requires challenges. They enjoy to always have a challenge to work on and also freedom to decide when to work on it, which connects to the description of mathematically highly able students as autonomous. Several students explicitly point out that they like that the work was not graded. Students differed in terms of working alone or with peers. The online enrichment offered mathematical challenges for most students, but some would have liked even more difficult tasks. Their involvement differed depending on other school work.

Does students' knowledge of mode correspond to what textbook tasks afford?

Karin Landtblom

Stockholm University

This presentation discusses the concept mode and includes findings from two studies. One study involved 130 students between the ages of 12 and 13 (school year 6). The students completed an online questionnaire in which they explained the meaning of the term mode and whether or not they thought mode was useful. According to the findings, pupils have a poor conceptual understanding of mode with little linkages to usability. The textbook analysis study involved seven Swedish textbook series. The tasks about mode were analysed based on input and output objects, and transformations. These categories allow for an analysis of the opportunities afforded to students to learn about the mode. Findings suggest that the textbooks do not afford enough possibilities to develop a deep understanding of the mode.

Designing a teacher guide to support teachers' planning, teaching and learning

Marcus Gustafsson

Karlstad University

This presentation reports on an ongoing project aiming to create more knowledge on how findings from research could be transformed into content of a teacher guide. If designed carefully, such a guide could potentially support teachers in "research-basing" their practice of planning and teaching, and also provide opportunities for teachers to learn by using the guide. Using a design research approach, a teacher guide will be tried out by teachers through an intervention in spring 2024. The initial design of the teacher guide will be guided by theoretically based design principles and build on findings from a systematic literature review on the teaching and learning of quadratic equations. How such a teacher guide potentially can support teachers' planning, teaching and learning is of interest for those designing teacher resources.

Fostering empowerment in STEM education: student-led curriculum development

Margaret Ann Bolick, Matthew Voigt and Kelly Lazar

Clemson University, United States

This presentation features an interdisciplinary exchange program between a Norwegian university and a United States university that repositions students as curriculum developers. Students are commonly viewed as passive learners in STEM classes, an orientation which fails to incorporate outside knowledge or recognize the life experiences and expertise of students. By engaging in this exchange program, five postsecondary students collaborated to lead a lesson study specializing in an interdisciplinary (geology and mathematics) place-based lesson set along the Savannah River Watershed in the Southeastern United States. This short communication presents an initial summary of how students engaged in the co-creation of the curriculum and insights into the lesson study implementation.

Diversity as a resource in mathematics teaching

Helena Grundén[1] and Helena Roos[2]

[1]Dalarna University, [2]Malmö University

This presentation is about diversity in mathematics classrooms, something all teachers need to relate to. A common way to meet diversity is Differentiated Instruction (DI). However, DI is questioned and might not be enough, hence, ideas from DI and previous research about inclusion in mathematics and planning have been used to develop a framework. The framework, Diversity Valued Instruction (DVI), emphasizes diversity as a resource in mathematics teaching. Results from a pilot study in which a group of teachers was presented with the framework show that the deficit perspective dominates the discussion and that adjustments are mainly about compensating for deficits. However, there is a glimpse of DVI, and teachers express care for students. The results contribute with important insights in the design of an upcoming study.

Students' talk about mathematical symbols

Ewa Bergqvist, Tomas Bergqvist, Fritjof Theens and Magnus Österholm

Umeå University

This presentation addresses natural language (for example Swedish) which is central to students' understanding of mathematical symbols. In the presented study, we focus on how students use natural language when talking about the meaning of mathematical symbols, and whether there are differences related to how familiar the symbols are. We look at two dimensions of how students talk, first on how explicit the students describe the meaning of a symbol. The second is what type of meaning is presented. This is work in progress, where we after several pilot studies have decided on methods for data collection and analysis. At the time for the conference, we plan to have carried out our main data collection and analysis. In the talk we will present preliminary results together with empirical data.

Investigating teachers' talk: challenges of classroom language data generation

Irina Johansson Carlén

Malmö University

This presentation is about discussing affordances and constrains of a tentative set of data generation methods regarding classroom language. This is a part of an early-stage PhD project, which investigates which types of registers that characterises mathematical classroom language within the early school years, focusing on the teacher's use of language registers. The presentation poses questions about data generation consisting of teachers' talk in whole class instruction that could be analysed quantitatively using corpus and SFL based analyses. Data will be generated and treated through three levels: initially, classroom lessons are recorded and transcribed; further on, transcription texts are treated through a series of suggested corpus linguistics software, aiming to produce a final data set fit for qualitative SFL analysis tools.

Mapping the landscape: a systematic review of implementation research in Nordic Studies in Mathematics Education

Iresha Gayani Ratnayake[1], Linda Marie Ahl[1], Uffe Thomas Jankvist[1,2] and Johan Prytz[1]

[1]Uppsala University, [2]Aarhus University

In this presentation we systematically examine implementation research reported in the journal Nordic Studies in Mathematics Education (NOMAD), focusing on mathematics education innovations implemented in Nordic countries over the past three decades. In particular, we sought to find answers to: What kind of innovations are reported in the papers related to implementation research in the Nordic countries in NOMAD. We identified four distinct kinds of innovations: a new teaching method, a new tool or resource for teaching and learning mathematics, a new method of learning mathematics, and curriculum reforms and textbook innovations. By identifying and categorizing different types of innovations, we contribute to understanding the research landscape in this field, especially in Nordic countries

Cognitive processing of mathematical symbols

Ewa Bergqvist, Bert Jonsson and Magnus Österholm

Umeå University

This presentation focuses on the use of symbols, which is crucial in the teaching and learning of mathematics, where it is essential to use natural language to describe their meaning. Language is sound based while symbols are primarily not but have also a visual dimension. Cognitive encoding of sound and visual input occurs in different components of working memory. This study shows that mathematical symbols are processed both through sound and visually, but that the sound-based processing plays a more significant role. This is also the case for unknown symbols, which is not in line with some previous research that suggests that new symbols should be processed more visually

Historical insights into statistical inference for K–12 education

Per Blomberg

Halmstad University and Karlstad University

This presentation outlines a study exploring the role of historical statistics in teaching practices to enhance students' understanding of statistical inference. Since the early 2000s, efforts have been made to integrate statistical inference into school curricula for a solid foundation in the subject. Using an educational design research approach at primary and secondary school levels, the study aims to develop didactical tools for teaching statistical inference. Preliminary findings from the initial phase suggests that historical perspectives offer didactic potential while highlighting substantial differences from current teaching practices.

Prospective primary teachers posing problems: some characteristics

Israel Garcia-Alonso[1], Diana Sosa-Martin[1], Mirela Vinerean[2], Karin Våge[2] and Yvonne Liljekvist[2]

[1]University of La Laguna, Spain
[2]Karlstad University

This presentation focuses on a preliminary study about the characteristics of the posed problems by prospective primary teachers (PPT). In this case, we analyse the plausibility, the reasonability and the mathematical structure of the 61 tasks on fractions that 21 students enrolled in the Primary Teacher degree have posed. The initial results imply that PPT are able to pose plausible and reasonable problems, but the variety of the mathematical structure selected could be improved by further development of the teaching practice on PTP. Educational design studies will be conducted in order to deeper understand problem posing as a tool for developing subject specific knowledge.

Cohesion and tension in task design: students working with multimodal tasks

Helena Johansson[1], Malin Norberg[2] and Magnus Österholm[1,2]

[1]Mid Sweden University, [2]Umeå University

This presentation focuses on mathematics tasks that challenge students, which have proven beneficial for learning. Tension (as an opposite to cohesion) between different modes (e.g., words, symbols) can create such challenge. By using cohesion and tension as concepts to analyse students' work on multimodal tasks, this study aims to identify characteristics in the task design that make students develop understanding of the mathematics concept in focus. Preliminary results show that a high degree of cohesion can direct the attention of students to the intended tension. Also, the tension in a task, although relevant, can be too challenging for students, thus ignored, and make students rely too much on previous knowledge.

Exploring the relationship between students' background, cognitive activation and mathematical knowledge development

Jimmy Karlsson

Karlstad University

This presentation reports on parts of a project with a focus on cognitive activation, which is a construct aiming at exploring how students potentially can engage in mathematical thinking to develop conceptual understanding. Previous research provides mixed results regarding the relationship between potential cognitive activation, students disposition, background and knowledge development. Preliminary results do not support a hypothesis that cognitive activation is more beneficial for students with more advantaged backgrounds. Further, the presentation aims to elicit and discuss aspects of cognitive activation, students' disposition, background and the relationship with knowledge development.

Teachers understanding of appropriate vocational mathematics knowing

Hanna Knutson

Gothenburg University

This presentation reports on tentative findings of a phenomenographic interview study, aiming to explore various ways of seeing what vocational mathematics knowing, appropriate for vocational students, entail. Semi-structured interviews, based on a set of geometry tasks related to a construction work context, were carried out with both vocational teachers and mathematics teachers. The analysis resulted in an outcome space of six categories of description, showing, for example, that 'appropriate vocational mathematics knowing' could be perceived as: Memorized facts and procedures; Context bound understanding; Vocationally relevant mathematical concepts; Integrated mathematical and vocational knowing. The categories are seen as complementary approaches, representing different facets of the phenomenon.

Voices from the field: implementing new education policy in classrooms and challenges with practice-based teacher education programs in India

Harita Raval

Stockholm University

This presentation explores challenges in implementing India's 2020 New Education Policy (NEP), focusing on the readiness of two teacher educators and their adaptation to its transformative goals. The study uses clinical interviews and Fairclough's discourse analysis to gain insights into the impact of NEP on practice-based teacher education, highlighting educators' struggles in aligning methodologies that NEP 2020 mandates. The perspectives of these two teacher educators reveal uncertainties, insufficiencies, and chaos, emphasizing the practical gaps between policy expectations and implementation. The presentation serves as a reflection on the challenges faced in Indian teacher education, underscoring the need for collective efforts to bridge gaps between international benchmarks and contextual dynamics.

The role of elicitation in formative assessment

Kristoffer Arvidsson

Umeå University and Umeå Mathematics Education Research Centre

A central component of formative assessment practices is the act of elicitation, where teachers create situations (e.g. through questions or tasks) that enable the gathering of information on students' current understanding and misconceptions. This short presentation will report on a case study, exploring the elicitation techniques used by one mathematics teacher when she responds to students seeking help during task-solving. Data is collected through audio recordings of lessons over the span of one year. Preliminary findings suggest that the teacher's feedback more often aligns with students' learning needs when the elicitation process yields sufficient information, i.e. enough information to enable the teacher to tailor the feedback to those specific needs. This indicates the importance of the elicitation process in formative assessment practices.

Tertiary mathematics students' foregrounds

Erik Adolfsson, Lisa Österling and Nhu Truong

Stockholm University

In this presentation, we bring the case of Chloe, a tertiary mathematics student, and her story about foregrounds, hopes and desires in mathematics. The presented study is part of a larger project , researching motives for studying mathematics, as well as aspirations for the future, among minoritized tertiary mathematics students. We use the construct of foregrounds to describe tensions between personal aspirations and collectively established expectations. From the five participating students, we present the story of Chloe. The story accounts for complex relations and tensions between a personal desire to pursue an interest in mathematics, and the collective expectations on entering mathematics studies for future careers. Such stories about diverse foregrounds increase our awareness of motives for entering tertiary mathematics.

Implementing a French teaching program in a Swedish grade 1 classroom

Anna-Lena Ekdahl and Klara Kerekes

Jönköping University

In this presentation we share experiences of having implemented a French teaching program in a Swedish grade 1 classroom for the duration of one school year. The French teaching program, Arithmetic and Comprehension at Elementary School (ACE) takes a structural and relational approach to teaching and learning numbers in early grades. The aim of this presentation is to describe the implementation process of the ACE-program in a Swedish context and to discuss how principles of variation theory can develop the program. Following the structure and the principle of the program we found that the program, with small adjustments, is possible to implement in a Swedish grade 1. We suggest that the program may be further developed by adding principles of variation theory.

Exploring Finnish and Swedish teachers' emerged classroom practice

Tuula Koljonen

Linköping University

This presentation reports on a case study that examines four Finnish and four Swedish primary school teachers' practices, utilizing the same original Finnish curriculum materials. Data consists of three video-recorded mathematics lessons, of which one per teacher is analysed in this presentation. The analyses uncover notable differences in classroom practices between the teachers in the two countries. The Finnish teachers utilize a more comprehensive range of questioning techniques, fostering active student participation, whereas the Swedish teachers deploy a more limited set of questions. This research contributes to the ongoing discussions about the nature and quality of instructional approaches in mathematics education from the implementation in two Nordic contexts.

Mediering via samtal vid formulering av en generell lösning till ett matematiskt problem

Hanna Fredriksdotter

Uppsala universitet

Presentationen fokuserar på utdrag ur mellanstadieelevers autentiska samtal om lösningar till matematiska problem, som exempel på hur formuleringen av en generell lösning kan medieras via samtal med klasskamrater. Forskning visar att gemensamt arbete är gynnsamt för elevers matematikutveckling; många studier har dock utförts som interventioner eller experiment, vilket innebär att det saknas kunskap om processer som kan uppstå under samarbete. I presentationen uppmärksammas särskilt hur eleverna delade med sig av information, byggde vidare på olika deluppgifters resultat, samt bekräftade lösningen genom tillämpning på resultat till redan lösta deluppgifter. Jag föreslår även hur innebörden av handlingarna kan omformuleras till instruktioner.

Lärares tal om matematiska situationer

Malin Norberg och Helena Vennberg

Umeå universitet

Den här presentationen beskriver en pågående studie om lärares arbete matematikklassrummet. Tillsammans med två verksamma lågstadielärare utforskar vi matematiska situationer i matematikundervisningen. Syftet med studien är att bidra med kunskap om vad som villkorar lärares möjligheter att utveckla tillgängliga matematiska situationer. Studiens datamaterial består av videoupptagningar och lärarnas reflektionsloggar. Preliminära resultat kommer att kunna presenteras på konferensen. Målsättningen är, förutom att uppnå studiens syfte, även att utveckla ett verktyg som kan utgöra stöd för lärares utveckling av matematikundervisning.

Om likhetstecknet som ”blir” eller ”är”

Robert Gunnarsson[1] och Emma Persson[1,2]

[1]Högskolan för Lärande och Kommunikation
[2]Handskerydsskolan, Nässjö kommun

I denna presentation diskuterar vi förklaringsmodeller för elevers förståelse av likhetstecknet. Matematikdidaktisk forskning har ofta beskrivit elevers förståelse för likhetstecknet som att de antingen har en operationell (dynamisk) förståelse eller en relationell (statisk) förståelse. Vi har under intervjuer med elever i lågstadiet sett att eleverna tycks pendla fram och tillbaka mellan operationell och relationell förståelse, båda tycks finnas parallellt. Vår slutsats är att det behövs fler förklaringsmodeller än statisk och dynamisk uppfattning för att beskriva elevers svårighet med likhetstecknet.

Elever och olika bedömningspraktiker i matematik

Maria Silwer

Malmö universitet

Denna presentation diskuterar en kommande studie med ett elevperspektiv på bedömning i matematik, i tidiga skolår. Resultatet av en pilotstudie visar att elever beskriver två olika diskurser i den matematikundervisning de möter i klassrummet; en där bedömning huvudsakligen sker i samband med prov och en där bedömning huvudsakligen sker i samband med problemlösning. Dessa två diskurser påverkar de roller elever beskriver att de kan inta som lärande. Diskurserna påverkar också hur elever använder sig av den feedback de får. Som doktorand avser jag fortsätta på pilotstudien och då jämföra hur elever och lärare beskriver bedömning i matematik, vilket jag i detta forum skulle vilja diskutera mitt övergripande upplägg av.