NYSESAM

A report on the NySeSam pilot project:
Curious Collaborative Observation of
Mathematics Teaching

Prepared by
Camilla Hellsten Østergaard
Julie Vangsøe Færch

July 2026

UNIVERSITY
COLLEGE
COPENHAGEN

c

KOLOFON

Published by Københavns Professionshøjskole at Open Books https://openbooks.kb.dk/kp/

Date of publication: july 2026

ISBN: 97887-85371-20-1

DOI: https://doi.org/10.7146/kp.660

Citation for published version (APA):

Østergaard, C.H., Færch, J.V. (2026). *NySeSam: a Report on the NySeSam Pilot Project: a Curious Collaborative Observation of Mathematics Teaching*. Københavns Professionshøjskole.

Supported by TrygFonden ID: 155145

1 Foreword

This report documents key aspects related to the implementation of the pilot project *NySeSam*. The report brings together the knowledge and experience gained during the project and provides the basis for a number of recommendations concerning how a future full-scale quantitative *SeSam* project should be designed and organised so that it can be conducted with a high level of quality and research validity.

The report is structured in accordance with the four phases of *didactic engineering* and complies with the requirements of CONSORT 2010 – Extension to Pilot Trials (Eldridge et al., 2016). In addition, the report addresses a number of issues identified in the pilot project's original application and subsequent revisions (Rasmussen & Færch, 2021; Østergaard & Færch, 2022), including reflections on the project's practical implementation, test development, methodological adjustments, and related considerations.

NySeSam is based on a structured model for teacher professional development, in which qualified supervision and video recorded classroom episodes form the basis for a systematic, dialogic development of teaching practices. Through a focused, evidence-informed process, the project examines the potential of video-based supervision to strengthen the teaching profession and to support a more coherent progression in teachers' continuing professional development.

With thanks to:

Participating schools, mathematics counsellors, teachers, and students

Statistical consultant: Svend Kreiner

Budget: Nicolai Kaadekilde Eickhoff

Teacher educators: Peter Stege Havsager, Peter Sars Schewitsch, Signe Gottschau Malm, Monica Wettendorff Pedersen, and Inge Marie Lyngsø Olsen

2 Resumé - Executive summary

NySeSam is a Danish pilot project designed to strengthen mathematics teaching in primary schools by developing and testing a practice-oriented model for teacher professional development. The project integrates three key components: video-based supervision, teacher collaboration, and structured problem-solving.

Participating mathematics teachers engaged in a structured professional development programme comprising workshops, collaborative lesson planning, implementation of structured problem-solving, and systematic supervision based on classroom video recordings. Students' mathematical progress was monitored using a specially developed testing tool.

Qualitative findings, derived from interviews with participating teachers, indicate significant professional growth. Teachers reported greater sensitivity to students' mathematical reasoning, increased reflection on their own educational practices, and an enhanced ability to connect theory and practice. The collaborative format—particularly the integration of peer discussion and video-based supervision—was consistently highlighted as a catalyst for deeper learning and increased instructional awareness.

Quantitative data from student tests suggest positive trends in learning outcomes, with some indication of increased mathematical engagement and reasoning among students who participated in the intervention. These initial findings provide a promising basis for future large-scale impact studies.

The project has also identified a set of practical and pedagogical principles critical to the success of such professional development initiatives. These include allocating sufficient time for reflection and change, organising teachers in collaborative teams, conducting one supervision cycle per month, and ensuring access to meaningful student assessment data. In addition, the project developed a detailed manual for video-based supervision, which offers a research-informed structure for guiding reflective teaching conversations and has potential for broader application in teacher education.

NySeSam demonstrates that it is possible to create a meaningful and sustainable link between teaching classroom practice and teacher learning. Through its structured yet adaptable design, the project contributes valuable insights into how mathematics teachers can be supported in developing structured problem-solving teaching practices. As such, NySeSam lays the groundwork for future scaling and provides a model for impactful, evidence-informed professional development in mathematics education.

Table of Contents

3 Introduction

3.1 Background

The ideas underpinning the NySeSam pilot project are based on a solid research foundation concerning effective teacher professional development, as well as on experiences from the earlier SeSam pilot project (*Vi SEr SAMmen på undervisning* [We observe teaching together]). These ideas are also grounded in a broadly anchored professional and political ambition to strengthen students' academic outcomes in the Danish public school system. This overarching ambition is expressed in the objective that all students should, to a greater extent, "become as competent as they can." It has, among other things, led to a political target that 80 per cent of students should be "good at reading and mathematics." Although these objectives remain in place, both national and international studies—including PISA (OECD, 2023) and TIMSS (Rûchenpöhler, 2024)—indicate that many students continue to find it difficult to realise their full potential, particularly in mathematics. There is therefore an undiminished need to develop and test new models that can support students' learning.

Educational research widely agrees that the teacher is the single most important factor influencing students' learning outcomes (Fullan & Hargreaves, 2012; 2016; Sowder, 2007). It is precisely against this backdrop that the NySeSam project was developed, with the ambition of enhancing the quality of mathematics teaching through targeted and structured video-based supervision of teachers, with students' learning as the ultimate goal. More concretely, NySeSam offers mathematics teachers an opportunity to systematically develop their teaching practice and their knowledge of children's learning of mathematics.

NySeSam examines the use of video-based supervision as a tool for systematic teacher professional development, supporting teachers in their work to develop high-quality, inquiry-based mathematics teaching.

The project draws on the international approaches of lesson study (e.g. Groves et al., 2016) and structured problem-solving (e.g. Asami-Johansson, 2021; Stigler & Hiebert, 1999). Both lesson study and inquiry-based approaches to mathematics teaching have strong research foundations and are already used in various local contexts in Denmark. However, they have not previously been examined systematically.

The ambition of the NySeSam pilot project has therefore been twofold. First, to develop and test a new format for video-based supervision that can be implemented in existing school practices. Second, to develop and test a subject-specific assessment tool capable of measuring students' development in mathematics. The project is ambitious—not least in a Danish context, where there is no strong testing culture and where national tests and the lower secondary school final exams constitute the only nationwide instruments for assessing students' academic achievement.

3.2 Purpose

The NySeSam pilot project examined the feasibility of video-supported supervision for the sustained professional development of mathematics teachers, in order to determine how a subsequent quantitative study of effects could be carried out. Supervision in this context is understood not as a one-way process, but as collegial dialogue in which each participant provides valuable professional support and is a source of insight for the other. The same applies to the role of "supervisor" in the context of SeSam and NySeSam, where the term does not connote a position of unilateral authority or one-sided expertise, but rather a reciprocal and inquiry-oriented relation in which participants contribute different forms of knowledge and each perspective constitutes a valuable professional resource for the other.

The main purpose has been to explore conditions related to building capacity, redesign of tests, and the design of the intervention, including workshops and video-based supervision. Within the project, a number of specific objectives were defined that needed to be achieved in order to fulfil this purpose.

Specific objectives

- Qualification of supervisors (building capacity of teacher educators for video-based supervision as knowledgeable others).
- Development of a SeSam intervention design, including
  - Workshops (introductory workshops and booster workshops)
    - Open lessons (joint observation of mathematics lessons and reflection on students' learning)
    - Structured problem-solving
  - Video-based supervision
    - Further development of supervision manuals that can be used reliably by different facilitators
- Development of a student test (a measurement instrument that can be adapted to the grade levels taught by the participating teachers)
- Testing of a wait-list intervention design

Overall, the project has generated specific knowledge about appropriate adjustments to the NySeSam format in preparation for a larger project.

3.3 Development of the protocol

This report is structured on the basis of didactic engineering, a research methodology within mathematics education (Artigue, 2014; Barquero & Bosch, 2015). Didactic engineering is a design-based research method aimed at investigating and understanding didactic phenomena—here, the NySeSam project—through a systematic process consisting of four phases.

1. **Preliminary analysis**: In this phase, a theoretical and didactic analysis of the content of the video supervision is conducted. The analysis is extended in light of experiences from the SeSam pilot project. The purpose is to identify prerequisites and potential challenges for subsequent phases.
2. **Design**: Based on the preliminary analysis, SeSam interventions and materials are designed to be implemented in controlled teaching contexts. This phase may involve an iterative

process in which design is adjusted in response to experiences gained along the way (e.g. experiences from SeSam and NySeSam).

3. **Implementation**: The designed SeSam interventions are tested in practice, with data being collected through student tests as well as through interviews in order to document the relevance of the SeSam intervention from a teacher perspective.
4. **A posteriori analysis and linking back to theory**: The collected test data are analysed in order to assess the effect of the SeSam intervention. The results are linked back to the theoretical framework in order to contribute to the further development of didactic theory and practice.

By structuring the report in accordance with these phases, a close alignment is ensured between theoretical considerations, practical design, empirical testing, and theoretical reflection. This supports an in-depth understanding of the didactic processes examined in the project.

4 Preliminary analysis

This part of the report constitutes the preliminary analysis, the purpose of which is to establish the theoretical and didactic foundation for the video-based supervision that forms the core of the intervention.

The analysis draws on existing research on teacher professional development, including the use of video as a means of promoting didactic awareness and subject-specific dialogue. In addition, knowledge from the earlier SeSam pilot project is incorporated. In our overall review, Appendix A, we also focus on lesson study, reflective conversation, and on how the process surrounding video-based supervision should be supported by a supervisor.

The analysis aims to identify the prerequisites that must be in place for video-based supervision to function as a meaningful practice that supports the continuing professional development of teachers. At the same time, the analysis identifies potential didactic, organisational, and relational challenges and barriers that need to be addressed.

4.1 Summary of research review: Potentials and challenges of video-based supervision

The present review, Appendix A, maps and analyses the existing empirical and research-based literature on the use of video-based supervision of mathematics teachers. The review places particular emphasis on the potentials and challenges inherent in mathematics teaching practice. The main results are organised thematically around four central analytical categories: (1) video-based supervision as a format of practice, (2) learning potentials, (3) the significance of didactic resources and structures, and (4) challenges and dilemmas.

First and foremost, the review shows that video-based supervision can function as a strong, practice-near format in which teachers (and pre-service teachers) are given the opportunity to analyse and reflect on concrete teaching situations. By retaining and zooming in on specific sequences of teaching, opportunities are created for deeper, more precise, and often shared reflection on didactic choices, students' participation, and subject-specific processes (e.g. Fox &

Poultney, 2020; Gee & Whaley, 2016; Tarantini, 2021). This differs significantly from traditional professional support formats, where reflection is often dependent on memory and retrospective accounts of events. Anchoring video in the actual teaching moments thus creates a more nuanced understanding of the complexity and dynamics of teaching.

The review further highlights that video-based supervision is particularly effective in supporting the development of subject-specific didactic awareness, through which teachers develop the competence to identify, interpret, and respond to significant events in the classroom. This applies both to students' subject-specific contributions and signs of understanding or misunderstanding, as well as to teachers' own teaching choices. Video recordings make it possible to pause, revisit, and renegotiate understandings of what actually happened in the teaching situation, and why. This promotes an analytical approach to the practice of teaching, in which teachers do not merely act, but learn to observe and understand teaching through a more professional lens.

An important point of the review is that the effect of video-based supervision depends to a large extent on the available didactic resources and structures within which the supervision takes place. Several studies indicate that it is not the video in itself, but rather the way it is used—including the type of reflective questions posed, the focus on the analysis, and the degree of support provided by a supervisor—that is decisive for learning outcomes (e.g. Gelfuso & Dennis, 2017; Tülüce & Çeçen, 2018). The use of focused tasks of observation and structured formats of conversation has proven to be particularly effective means of supporting reflection and professional dialogue. In addition, the importance of the supervisor's role—the *knowledgeable other*—is highlighted, including their ability to facilitate inquiry-based, non-evaluative conversations in which participants feel safe to share uncertainties and mistakes.

Despite the many potentials, the review also identifies a number of significant challenges and dilemmas associated with the use of video-based supervision (e.g. Barth-Cohen et al., 2018; Koh, 2015; Tülüce & Çeçen, 2018). Several studies describe how video-based supervision can make teachers (and pre-service teachers) feel vulnerable when being filmed and exposed as objects of observation by others. This sense of vulnerability can inhibit reflection and openness in supervision, particularly if an evaluative or controlling atmosphere emerges. In addition, the use of video requires a certain level of technical infrastructure, time to view and analyse recordings, as well as competencies on the part of the supervisor to pose relevant questions and guide the professional and subject-specific focus of the conversation. Another central dilemma is the risk of an excessive focus on the teacher as the primary point of attention in the video, rather than on the students' learning and participation—which may shift attention away from the primary purpose of the supervision.

Finally, the review shows that video-based supervision holds significant potential for connecting theory and practice (e.g. Walkoe et al., 2020), especially when combined with clear subject-specific didactic goals and an inquiry-based approach. At the same time, it is emphasised that successful implementation presupposes clear structuring, competent supervisors, and a culture in which video analysis is viewed as a learning resource rather than an assessment instrument. The review thus points not only to the fact that video can make a difference—but to how and under what circumstances it actually does so.

4.2 Experiences and recommendations from the SeSam pilot project

The experiences from the earlier SeSam pilot project (2020) form an important foundation for the development of NySeSam. The SeSam project was structured around two different interventions: SeSam-LS (Lesson Study) and SeSam-CLASS, each of which tested different forms of video-supported supervision with the aim of strengthening teachers' professional development in primary and lower secondary school.

4.2.1 SeSam-LS: Video and subject-specific professional support in the context of lesson study

SeSam-LS was based on central principles from Japanese lesson study culture. Within this model, teaching was planned in close collaboration between the teacher and a supervisor, after which selected lessons were conducted, video-recorded, and subsequently analyses in online discussions based on the recordings. The focus of discussions was centred on three main areas:

a) the teacher's mediation of students' interaction with mathematical content,
b) students' subject-specific learning, and
c) the teacher's enactment of the planned lesson.

Between each video-based discussion, lesson plans were shaped and discussed, creating continuity and progression in the collaboration. The discussions took place at intervals of approximately two weeks, and recordings of lessons were conducted approximately once every four weeks. In SeSam-LS, supervisors were referred to as *knowledgeable others*, reflecting their role as subject-specific supervisor with specialised knowledge of mathematics, mathematics teaching, mathematics learning, and the didactics of mathematics.

4.2.2 SeSam-CLASS: Observation of general teaching qualities

The second intervention, SeSam-CLASS, was based on the American CLASS system (Classroom Assessment Scoring System). In this intervention, the participating teachers' mathematics lessons were observed and analysed on the basis of three overarching domains: *Emotional Support*, *Classroom Organization,* and *Instructional Support*. The observations were conducted by certified CLASS observers from University College Copenhagen, who used 15 minute video recordings as the data basis. The reflective discussions were based on ten so-called *dimensions* related to the three domains, and the discussions were occasionally supported by written prompts prepared by the CLASS observer. The frequency of contact was approximately every other week.

4.2.3 Common features and learning potentials

Across the two interventions, the SeSam system essentially consisted of three components:

1. systematic video documentation of teaching,
2. sharing of video recordings between teacher and supervisor, and
3. qualified subject-specific professional conversations about teaching.

Both interventions aimed to create conditions that enabled teachers to develop their teaching practice in dialogue with an external subject specialist.

The pilot project showed that teachers' preferences for intervention formats varied. Several schools explicitly expressed interest in participating in SeSam-LS, while a smaller number showed particular interest in the CLASS format.

4.2.4 Adjustments in NySeSam: Focus on subject specificity and dialogue

NySeSam carries forward the ambitions of the original SeSam project (application to TrygFonden).

In the new intervention, we have chosen to omit the CLASS format, a decision informed by both research-based considerations and concrete feedback from teachers and teacher educators. Several participants indicated that the evaluative character of the CLASS observation scheme was experienced as inhibiting subject-specific dialogue. The scheme focused on general didactic aspects such as classroom management and relationships, while the mathematical content of the teaching was only included to a limited extent. In addition, the intervention lacked a structure for working with lesson plans, which in SeSam-LS was highlighted precisely as a strength—since discussions of the planning of upcoming lessons proved to enhance both the sequencing of teaching activities and reflection.

4.2.5 Paradidactic structure and supervision in NySeSam

NySeSam builds on the structure characteristic of lesson study. This means that teachers and their supervisors from the university college collaborate and engage in discussions about lesson plans prior to teaching—and subsequently about completed lessons, drawing on video recordings. This structure was chosen on the basis of earlier experiences from SeSam and is underpinned by research-based knowledge on video-supported professional development, as presented in the accompanying literature review (Appendix A).

5 Design

Based on the preceding preliminary analysis, this section presents the design of video-based supervision in the SeSam intervention and the resources that form the basis for testing the intervention.

The design aims not only to establish a practically applicable model for video-based supervision, but also to function as a didactic experiment that enables systematic analysis of the SeSam intervention.

5.1 Design and programme theory for NySeSam

NySeSam is based on the programme theory shown in Figure 1 (Rogers, 2008).

The model provides a clear overview of the intervention and draws attention to contextual factors and assumptions about the mechanisms of change.

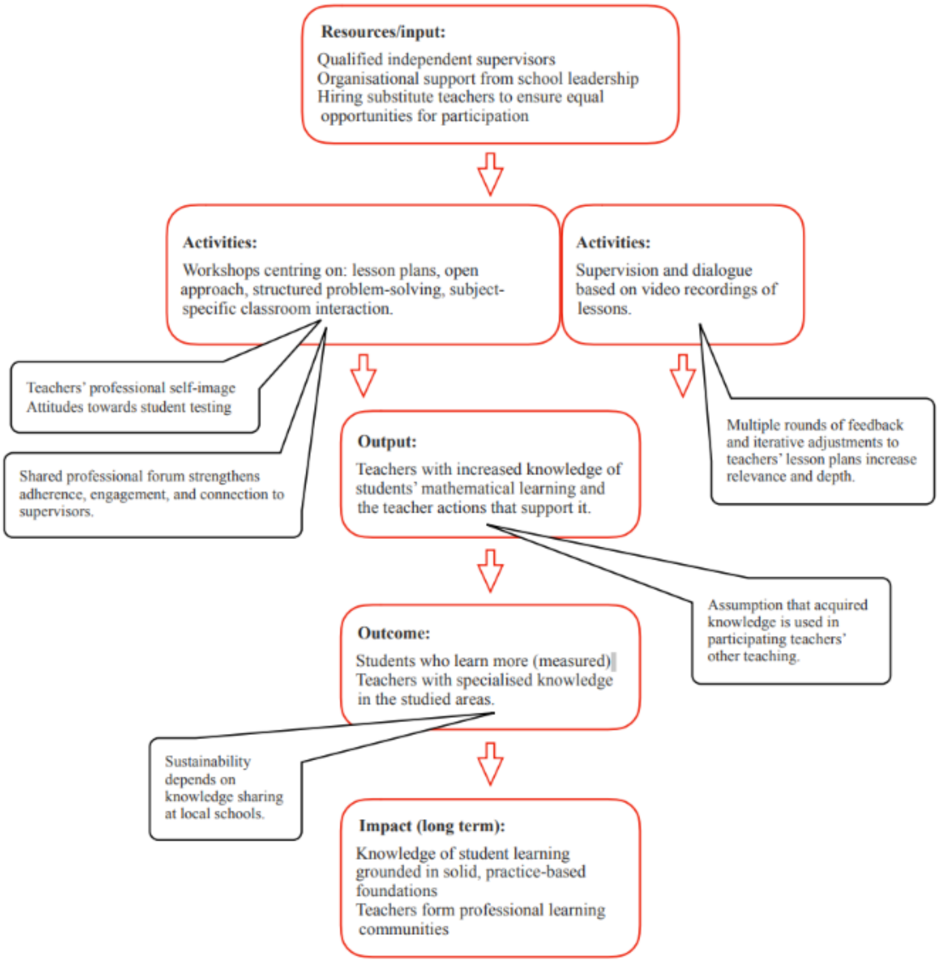


Figure 1. Programme theory for NySeSam

5.1.1 Intervention design

In NySeSam, we tested a wait-list design, illustrated in Figure 2, in which all participating classes served as their own control group. There are several reasons why this design was considered desirable to test in a project such as this.

First and foremost, the focus is not on a specific grade level, but rather on the development of teachers' competencies more generally. Students' age is therefore of secondary importance. At the same time, this should also make it easier in practice to recruit teachers, as there are no constraints related to the grade level at which the teacher teachers. In this pilot project, the delimitation to grades 1–6 was due to the fact that the *MatematikProfilen* materials for grades 7-9 were not ready at the start of the pilot project. The test materials for grades 7–8, in the form of separate *MatematikProfilen* editions, will be ready for the 2025/26 school year. As the aim has been to include teachers broadly across grade levels, we consider the wait-list design to be advantageous, as it ensures comparability in group size and socio-economic background variables among the participating students.

In addition, the design makes it possible to initiate the SeSam interventions on an ongoing basis, which makes the project more flexible for the participating teachers. In NySeSam, the rolling initiation has meant that some schools started in August (period 0, Figure 2), while other schools were initiated in June (period 1, Figure 2). As a result, the control period for schools in period 1 included a summer holiday.

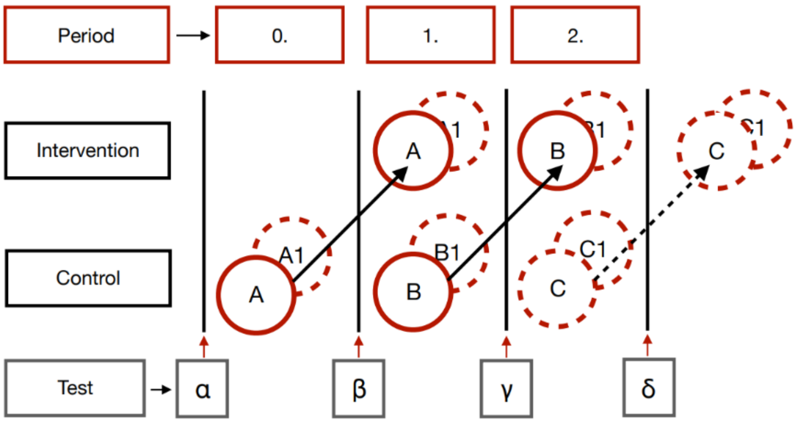


Figure 2. Elements of the intervention

5.1.2 SeSam intervention

The SeSam intervention consists of the following (see also the outlines for teachers, *information and cadence*, Appendix C):

- An initial workshop:
  - Practical information about the NySeSam project
  - Technical equipment, video recordings, and how videos are uploaded to Teams
  - Structured problem-solving
  - Open lesson I:
    - Review of the lesson plan and format
    - Joint observation of a structured problem-solving lesson
    - Joint reflection on students' learning in the lesson in question
- 2 x 2 individual video-based supervision session
  - Two rounds, each consisting of:
    - Supervision based on teacher-produced lesson plan
    - Supervision based on video from the teacher's enacted lesson
- Boost workshop
  - What constitutes good inquiry-based problems?
  - Open lesson II (preferably conducted by one of the participating teachers):
    - Review of the lesson plan
    - Joint observation of a structured problem-solving lesson
    - Joint reflection on students' learning in the lesson in question
    - Focus on how to make high-quality observations
    - Focus on written work in mathematics
- 2 x 2 individual video-based supervision sessions
  - Two rounds, each consisting of:

- Supervision based on teacher-produced lesson plan
- Supervision based on video from the teacher's enacted lesson

5.1.3 Structured problem-solving

Structured problem-solving is an inquiry-based approach developed by teachers and originating in Japan (Asami-Johanson, 2021). The purpose of structured problem-solving is first to arouse students' curiosity and subsequently to support their development of new knowledge. The progression of a structured problem-solving lesson can be described in the following phases (Asami-Johanson, 2021; Rasmussen & Isoda, 2019; Souma, 2017):

- The teacher presents the problem: An initial problem is posed with the aim of engaging the students.
- The students engage in understanding the problem.
- The students propose hypotheses or suggest solution methods.
- The students reformulate the derived problem in relation to the original problem.
- The students engage with the problem: the mathematical inquiry phase. The teacher observes the students' strategies to later be able to orchestrate a comparison of students' solutions.
- Whole-class discussion: The aim is for students to improve their methods and challenge their thinking by drawing on mathematical reasoning. This requires careful didactic orchestration and staging on the part of the teacher.
- The teacher invites students to compare, negotiate, refine, and develop new perspectives through elaborative questions about students' solutions.
- The students are encouraged to listen to one another's ideas and assess strengths and weaknesses, simultaneously creating connections between solutions.
- Teacher and students jointly summarise their learning with a view to elevating the students' mathematical understanding to a higher level.

We proceed from the hypothesis that structured problem-solving can contribute to expanding students' mathematical knowledge.

Building capacity – supervisor

Building capacity and developing the supervisor (knowledgeable other) has been given high priority in the NySeSam pilot project. Existing research (Appendix A) shows how the supervisor plays a central and important role in professional dialogue and in teachers' development by

- sharing different and alternative perspectives on teaching
- make connections between theory and practice visible
- supporting teachers in reflecting on their teaching and students' learning

As there is no formal education programme for supervisors in a Danish context, supervisors from University College Copenhagen (KP) have participated in a specially designed programme at Hokkaido University of Education, Kushiro in Japan. Here, the KP supervisors participated in observations and reflections within already established supervision environments at Kushiro Fuzoku School (a school affiliated with Hokkaido University of Education, Kushiro).

During the stay in Kushiro, KP supervisors

- participated in 24 structured problem-solving lessons,
- observed how the school's experienced teachers acted as knowledgeable others (supervisors),
- discussed the lessons in question with experienced teachers who, in lesson study contexts, act as knowledgeable others,
- discussed and developed a dialogue guide for NySeSam supervisors (Appendix E), thereby ensuring consistency in the supervision dialogue,
- participated in virtual lesson studies.

Through the stay in Kushiro, KP supervisors have gained valuable experience in concrete ways of developing the role of the supervisor in a lesson study context. Through observations, participation in planning meetings, and subsequent reflection, the participants gained insight into how Japanese educators systematically build and apply specialised knowledge to support teacher professional development. This has given rise to reflections on which competencies and forms of practice are required to fulfil the role of the supervisor in a meaningful and qualified manner, and how these can be strengthened and developed in a Danish context.

5.2 Method

5.2.1 Practical arrangements

5.2.1.1 Recruitment of schools

In NySeSam, the original idea was to recruit 24 teachers, who would then be divided into four groups and each be assigned their own SeSam supervisors.

In the process, we sent unsolicited letters to teachers and schools. In practice, this resulted in the recruitment of 22 teachers, who teach mathematics in 31 classes across three schools. In total, 745 students participated. The distribution of students across the participating grade levels is shown in Table 1.

Table 1. Total number of participating students

Grade level	1	2	3	4	5	6
School 1	53	60	61	65	50	66
School 2	44	70	50	28	27	
School 3	51	71			49	
Total	148	201	111	93	126	66

Recruitment has been a challenge. This is despite the fact that schools had substitute teacher costs covered and that teachers participated free of charge in teacher professional development.

5.2.1.2 Experiences in relation to future recruitment of schools

In the final phases of the NySeSam project, we have drawn on the experiences that *SYstematisk KLassekammerathjælp* (SYKL, supported by TrygFonden and Novo Nordisk Fonden) has developed through its involvement of more than 135 schools across Denmark in an RCT study.

A future recruitment strategy, inspired by SYKL, includes:

1) A shared Excel spreadsheet containing the approximately 1,400 Danish public schools. This is intended to function as an overview of schools, their allocation, and the status of contact with each school.
2) Allocation of schools across the country among project participants from University College Copenhagen. Direct written and telephone contact should be made with school leaders, mathematics counsellors, and municipal consultants.
3) Organising and hosting regular information webinars.
4) Professionally produced videos presenting the project.
5) Social media posts.
6) Presenting the project at national conferences for mathematics counsellors and municipal mathematics consultants.
7) Presenting the project within municipal networks for mathematics counsellors.
8) Presenting the project at municipal network meetings for school leaders, possibly facilitated through municipal mathematics consultants.
9) Producing articles about the project for MATEMATIK (approximately 4,000 teachers) and Folkeskolen.dk.
10) Text for the website of NCUM (National Centre for the Development of Mathematics Education).
11) A dedicated NySeSam website for registration and information.
12) General outreach to the press.

5.2.1.3 Information materials

As part of the implementation of the NySeSam pilot project, we have developed and distributed targeted information materials to both schools and parents (Appendix G). The purpose has been to ensure transparency, support, and informed participation from all involved parties. The materials consisted of a general introduction to the project's purpose, content, and timeframe, as well as a description of the activities in which teachers and students would participate—including structured problem-solving and video recordings of teaching.

In addition, a separate information letter has been prepared in connection with the data collection (Appendix F), which clearly sets out how data are collected, processed and stored in accordance with applicable regulations on data protection and ethical approval. The letter made it possible for parents to make an informed decision. It also emphasised that participation was voluntary and that all information would be treated confidentially.

The information effort has been important in creating a sense of security and trust among school leadership, teachers, parents, and students, and it has at the same time ensured that all involved have had a clear understanding of the project's intentions and methods. The materials may serve as a model for future communication efforts in a future quantitative study.

5.2.1.4 Formulation and further development of a manual for video-based supervision

Throughout the project period, we have systematically developed and refined a manual for video-based supervision, which has served as a central tool in the NySeSam project (Appendix E). The purpose of the manual has been to support a professional and consistent approach to observation, analysis, and dialogue about teaching based on video recordings. It has primarily functioned as a support tool for the supervisor. The point of departure for the manual has been to create structure in the video-based supervisions, so that the focus is kept on the content of the teaching and on students' learning rather than on the teacher. The manual specifies, among other things, guidelines for how video excerpts are selected, which aspects of the teaching are brought into focus (e.g. student statements, teacher questions, reasoning and arguments), and how the conversation about the excerpt can be facilitated through open, inquiry-based questions. The purpose has been to create an inquiry-based and non-evaluative form of dialogue in which teachers can safely share, analyse, and learn from concrete teaching episodes.

Experiences from the NySeSam pilot project have led to several adjustments and clarifications in the manual, including the structuring of the conversations and a clearer articulation of the supervisor's role as a facilitator rather than an evaluator. The manual now constitutes a well-founded and tested tool that can be used broadly in both initial teacher education and teacher professional development, and that may potentially contribute to enhancing the quality of collegial professional dialogue and support through video.

5.2.2 Quantitative data

5.2.2.1 Changes in the test setup from SeSam to NySeSam

The National Tests in mathematics were previously used in the SeSam project, and although the new mathematics National Transition Tests are administered in grades 2, 4,6, 7, and 8[1], this would still limit our ability to include all grade levels in a future, larger project. Moreover, the National Transition Tests do not allow for the implementation of the desired wait-list design. In addition, the National Transition Tests are not criterion-based and therefore cannot provide information about students' mathematical knowledge, these tests were not selected as part of the test setup in NySeSam.

In the NySeSam application, we described a test setup involving self-scoring tests developed and administered on the online platform *Matematikfessor*. When NySeSam was launched, however, *Matematikfessor* had been acquired by Alinea, and it was no longer possible for us to adapt and develop the desired types of tasks. We were able to use the existing multiple-choice or numerical tasks, but we were not able to use tasks that assessed students beyond the highly skills-oriented tasks already available on *Matematikfessor*.

[1] Folkeskolens Nationale Overgangstest | Børne– og Undervisningsministeriet [National Transition Tests]

We therefore searched for an alternative assessment material that tested more broadly. In this search, we came across Gyldendal's *MatematikProfilen* (MP), which is a relatively comprehensive test material that, at each grade level, consists of three parts. The total test time is approximately three hours. The first part consists of dichotomous multiple-choice items. The second and third parts consist of polytomous items which, in addition to testing knowledge of different mathematical content areas, also assess students' mathematical competencies, such as reasoning competency, representational competency, and symbol manipulation competency. For these reasons, they provide broad coverage in relation to the current Danish National Curriculum (*Fælles Mål*).

In the NySeSam pilot project, the purpose is to test whether an assessment material such as MP can measure development by comparing student scores before and after the SeSam intervention is implemented. This presupposes that development can be assessed in such a way that students' academic level is measured at different points in time on one and the same scale, without responses to tasks at one point in time being influenced by responses to tasks at earlier points in time. In Appendix I, we describe how this can be achieved by means of three subtests consisting of items from the MP tests.

5.2.2.2 *Pedagogical tests*

For pedagogical tests to meet the requirements of precision, they often consist of many different tasks which, in addition to placing limitations on their practical use, can also be demanding and challenging for students due to the amount of time they are required to spend on the tasks. Situations may therefore arise in which it would be convenient to reduce such tests to a smaller number of short tests, without compromising anything other than the precision of the test results.

The use of MP in NySeSam illustrates such a situation. In addition to reducing the number of tasks students are required to respond to on each occasion, it provides a material that measures students' development over a given period by examining changes in test scores, without requiring students to respond to the same task multiple times.

At each grade level, the MP tests consist of 32 dichotomous items, with responses recorded as either correct or incorrect; 16 trichotomous items, with scores of 0, 1, or 2; and 6 polytomous items with scores ranging from 0 to 3. The total score across the full material ranges from 0 to 82.

The tasks fit a so-called log-linear Rasch models, in which—unlike conventional Rasch models—local dependencies between items is permitted.

One of the major advantages of using a test material such as MP is that extensive piloting and analyses of the tests have already been conducted, and this work can be drawn on in various ways. On the one hand, we have the opportunity to use Gyldendal's MP data to verify that MP's Rasch models also fit data from NySeSam. On the other hand, from a quantitative psychometric perspective, we know that MP meets requirements of validity and objectivity due to the fit of both previous and current data to the Rasch model. That the measurements are precise and free from systematic bias can be seen from the tables in Appendix I, where the standard error of measurement (SEM) of the Rasch model's person parameter (WML estimate) for the vast majority of students lies well below 0.30 in our customised version of MP. In the original MP test material, the SEM is even lower than here. By comparison, the National Tests were intended to reach an SEM level of 0.30 or below but rarely did so.

A decisive advantage of using MP, compared with many other assessment materials on the market, is that the material is criterion-based[2]. The fact that the material is criterion-based means that, in addition to being able to determine whether there is an effect of an intervention—as measured by changes in test scores before and after the intervention—we can also articulate in what ways the participating students have become more proficient. MP uses five categories: *Started* (Begyndt), *In progress* (I gang), *Well under way* (Godt i gang), *Beyond expectations* (Længere end forventet), and *Far beyond expectations* (Meget længere end forventet). At each grade level, the profiles provide a description of which mathematical content areas students in the different categories have a strong grasp of, and which areas still leave room for improvement. The criterion-based nature of the material therefore also makes it highly applicable for the participating teachers. An example of the feedback provided to the participating teachers in NySeSam can be found in section 7.1.1.

Overall, we therefore consider MP to meet all requirements for pedagogical tests in which test results must be valid, objective, precise, free from systematic bias, interpretable, practically applicable in classrooms, and useful.

5.2.2.3 *Partitioning of test material*

To move from the full MP material to a material divided into three subscales, the construction and testing were carried out in two steps. In the first step, a number of items are selected that are locally independent of the remaining items in the full test. If the full set of items fits the psychometric measurement model, the same will apply to the selected items, and it is therefore possible to use the item parameter estimates to an estimate of the person parameter. The estimate of the subscale's estimate of proficiency is calibrated to the scale of the full test, but confidence with respect to potential bias will be lower than for the proficiency measure based on the full scale, because the subscale contains fewer items.

There is a one-to-one correspondence between the total number of points on a test that fits a Rasch model and the estimate of proficiency on the Rasch model's logit scale. One can argue that the Rasch model's logit scale is a so-called interval scale, but the scale is abstract, and most users of the test results therefore prefer to use the total number of points rather than the logit values. To calibrate the total number of points on the subscale to the total number of points on the full test, we have therefore included calculations of the expected total score on the full test based on estimates of the logit values from the subset. Calculations for the three subscales can be found in Appendix I for each grade level. The weighted maximum likelihood estimate (WML) is the preferred estimator for person parameters in the Rasch model. In addition to WML, the tables also present the maximum likelihood estimate which is used to calculate the expected number of points on the items that are not included in the subtest.

[2] The profiles for each grade level can be found on the MatematikProfilen website, which all teachers can access with a UNI login via: Resultater | Matematikprofilen | Gyldendal

The second step in the testing of the subscales consists of analyses of the collected data using the subscales, examining whether the Rasch model for the full test fits the test results when collected at the subscale level. In addition to the expectation that this will be the case, it is also expected that some items may function slightly differently in the full MP test compared with our partitioning. This is because responses to one item cannot be assumed to be independent of responses to another item, which means that the collection of data for the subscales differs across the two test setups.

Should it be the case that some items behave differently in the two test setups, the estimates of the person parameter will be measured on the same scale as in the original MP. It is therefore possible to convert the number of points in NySeSam's subscales into expected results, as if they had been collected in the same way as in the original administration of MP. The result can then be converted into an expected total score on the full MP scale. Appendix J shows how the NySeSam results from the three subscales are to be converted into results, as if the items in the subscales had been answered together with the remaining items in the full MP materials. As expected, only limited adjustments are required.

In our customised version, the total score for MP1-3 does not reach 82 as is the case in the original MP material. This is because one third of the multiple-choice items at grades 1-3 do not test the same dimension as the remaining items in MP. As these items are defined on a different scale than the rest of the items, they are not included in the partitioning, which is why subtests 1-3 at grades 1-3 only reach a maximum of 69 points.

5.2.3 Qualitative data

5.2.3.1 Interviews with teachers

In its basic structure, the NySeSam project is designed as a quantitative study with a focus on measuring effects. However, in order to supplement the quantitative data and gain a deeper understanding of participants' experiences and perspectives, we have also included qualitative elements in the form of semi-structured interviews. Semi-structured interviews are a qualitative method in which the interviewer follows an overarching interview guide with key themes, while at the same time allowing participants to elaborate on and nuance their responses. This approach makes it possible to collect rich, context-sensitive data, while also ensuring a certain degree of comparability across interviewees.

In connection with interviewing participating teachers and supervisors, we have developed an interview guide (Appendix D) which serves as a tool to ensure that key themes and questions are addressed across interviews. The guide contains a set of overarching questions and themes related to participants' experiences with the project, their perceptions of the collaboration, and their reflections on the development of their teaching and their own learning. At the same time, the interview guide is designed with a certain degree of openness, allowing for follow-up on relevant points introduced by the participants themselves. This is in line with the semi-structured approach where the interview is guided both by the research interest and by the participants' own perspectives and experiences.

6 Implementation

The implementation of NySeSam has been a central focal point in which ambitions and design of the pilot project have been tested in practice. To realise this, the implementation was carefully planned, but also flexible and adaptable, making it possible to respond to local needs, practical conditions, and participants' experiences along the way. Throughout the implementation, teaching plans and activities, test designs, and videos of practice have been developed and adapted on an ongoing basis. It has been essential to ensure that the participating teachers developed a meaningful sense of ownership of the project, and the implementation has therefore focused on both structural support and professional engagement with subject matter and content.

In the following, we describe how the implementation of NySeSam has been carried out in practice, including the practical administration of *MatematikProfilen* (MP). In addition, teachers' voices are foregrounded, as their reflections and experiences shed light on how the project has been experienced and taken up in practice. Through these perspectives, we gain insight not only into implementation as a set of planned activities, but also into implementation as a dynamic and lived process in practice.

6.1 The practical administration of MP

At School 1, subtests A, B, and C were administered in September, December, and March of the 2023/2024 school year, while the subtests at School 2 and 3 were administered in June, August, and November 2024. An annual timeline illustrating test administrations and interventions at the individual schools is shown in Figure 3.

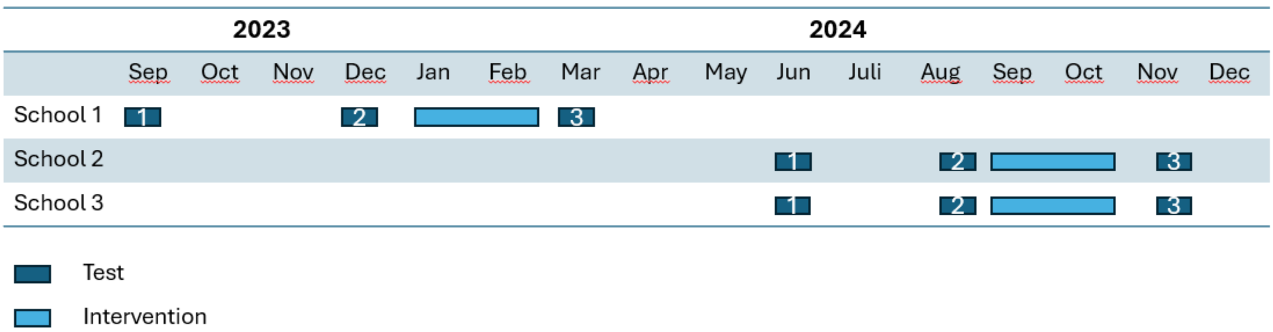


Figure 3. Timeline: NySeSam

For each subtest, a manual for administration has been developed which schools are required to use to ensure consistency across classes. These manuals can be found in Appendix H.

Of the 31 classes, three withdrew during the course of the project. In addition, students who did not complete all three subtests were excluded prior to the analysis, as we aim to compare development during the initial months without the SeSam intervention with the development during the months with the SeSam intervention. This exclusion resulted in 512 participating students, as shown in Table 2:

Table 2. Number of students participating in all three subtests

		Grade level					
Test occasions	School	1	2	3	4	5	6
September – December – March	1	39	49	48	38	42	57
June – August – November	2	21	44	21	21	22	
	3	24	42			44	
	Total	84	135	69	59	108	57

6.2 Interviews: Teacher voices in the implementation of NySeSam

As a central component of the NySeSam pilot project, qualitative interviews have been conducted with participating teachers (interview guide, Appendix D). The purpose has been to gain insight into teachers' experiences, understandings, and reflections in connection with their participation in the project. This is both in relation to their own professional development and in relation to their teaching and perception of students' learning. The interviews have provided a valuable opportunity to understand how teachers have experienced participating in a new and structured development process in which mathematics teaching, shared video-based collegial dialogue, and an inquiry-based approach to practice have been central.

The conversations have not merely functioned as an evaluative activity but have also opened space for teachers' own considerations, as well as clarifications related to subject matter and content. Reflections centre, among other things, on what it has been like to make one's own teaching the object of video-based collegial dialogue and support, to participate in a structured framework for video recording and supervision, and on what it has required and afforded to engage in inquiry-based mathematics in their own classroom.

The interviews thus offer insights into teachers' perspectives on professional and subject-related development. Statements from teachers express enthusiasm, resistance, moments of realisation, and curiosity. In what follows, selected quotations from three of the interviewed teachers—here referred to as A, B, and C—are presented. Taken together, they provide a nuanced picture of how the project has affected them in subject-related and personal terms.

6.3 Teacher voices

Teacher A describes participation in the project as an eye-opener: "It was incredibly inspiring […] it was really enjoyable to suddenly be presented with a mathematical problem that you yourself became absorbed and immersed in. And going through this ourselves is what really gave us a different kind of insight […] This is the kind of experience I want to share with my students."

Teacher A experiences that participation has led to development in their work as a teacher: "We almost had to be forced into it […] honestly, I have so many colleagues who could also really use being shaken up, just like I needed it. […] When we take the focus away from ourselves and instead look at the concrete situation, we become better at what we do."

The project has also had an impact on a personal level: "If I can do it in front of all my colleagues […] then it is something I can become better at myself. […] It requires extra preparation […] I had to change some set ways of thinking and be much more conscious of […] choice of words, sequencing […] It was demanding. But it was inspiring. It was fun."

Teacher B emphasises the rare experience of themselves receiving substantive subject-related input: "First of all, it is really nice when you are not used to being the one who gets input yourself."

Collaboration with colleagues around lesson design is highlighted as meaningful: "We started by asking: what is it, actually, that we want to investigate? […] We found something we could do at both grade levels, in classes one year apart."

However, participation has also been demanding: "It felt like an exam every time. […] I felt that what I needed to practice was staying in the conversation."

At the same time, the teacher expresses reflection on the balance between inquiry and students' needs: "I also think the children reach a point where they've had enough […] can't we just be allowed to sit and work together on a task? "

Teacher C emphasises the value of collegial dialogue and support: "When you've had some thoughts about something, it's really helpful to have a second pair of eyes look at it as well." At the same time, the teacher highlights how the systematic design challenges their usual practice: "I usually tend […] to be very structured […] but it's actually really great to have some variation in teaching and learning."

At the same time, a relevant question is raised: "But does it actually work? I mean, is it, for example, a better way of learning?"

7 A posteriori analysis

Following the conclusion of NySeSam, an a posteriori analysis has been conducted with the aim of examining which insights and results can be derived from the pilot project's data. The purpose of the analysis has been to shed light on the extent to which, and the ways in which, the project's core activities—including the use of inquiry-based mathematics teaching, structured video-based supervision, and collegial dialogue and support—have contributed to the development of both students' mathematical progression and teachers' professional practice. The analysis draws on two primary data sources: first, quantitative test data from students who have participated in the project's activities, and second, qualitative interviews with the teachers involved. The tests have made it possible to track students' development over time and to identify indications of increased understanding and reasoning in mathematics. The teacher interviews have provided deeper insight into how the teachers have experienced their own subject-related development and the changes they have made to their teaching practice as a result of participating in NySeSam.

7.1 Analysis of student progression in NySeSam

Overall, the pilot project's test results show a clear improvement among students, particularly in the lower grade levels. In the following, grade 1 is used as an example. The average change in test score from Part A to Part B is 2.6 points, while the change from Part B to Part C is as much as 9.8 points. The marked improvement between Part B and Part C indicates that the intervention has had a positive effect, and that the test at the same time functions as intended, as a greater shift in students' level of mathematical attainment is observed during the intervention period than under ordinary teaching conditions.

Statistically, the analysis shows a highly significant difference in test scores from Part B to Part C (one-sided p-value < 0.001), whereas only a weakly significant different is observed from Part A to Part B (one-sided p-value $= 0.039$). This difference supports the hypothesis that the intervention has contributed to increased learning. Across all participating schools at grade 1 level, a statistically significant improvement from Part B to Part C is observed, albeit with varying strength.

The change in score itself does not, however, say anything about *how* students have become more proficient. As the test is criterion-based, it has been possible to supplement the analysis with analyses based on the five response categories defined in the original test material: *Started*, *In progress*, *Well under way*, *Beyond expectations*, and *Far beyond expectations*. Shifts in these categories are observed from Part A to Part B and from Part B to Part C. From Part A to Part B, the analysis shows a statistically significant difference in the distribution ($p = 0.005$) with a moderate association (gamma coefficient = 0.26), whereas the difference from Part B to Part C is stronger ($p < 0.001$, gamma coefficient = 0.66). This shift is also reflected in the proportion of students who are placed in the upper categories (*Beyond expectations* or *Far beyond expectations*), which increases from approximately 20% in Part B to more than 65% in Part C (Appendix I). By comparison, the increase from Part A to Part B was only 5 percentage points. This development indicates that students have made substantial progress during the intervention period.

For grade 5, a different picture emerges. Here, it has not been possible to demonstrate any significant effect of the intervention, likely due to the small sample size, which has generally made it difficult to identity effects for grades 3–6. A surprising finding for grade 5 is that the average test score is lower in Part B than in Part A, a difference that is highly significant ($p < 0.001$) but not meaningful, as it would otherwise imply that students have become less proficient. This suggests that the administration of Part A at some schools was not conducted under the same conditions as the subsequent tests. At several schools, students are observed to perform unusually well in Part A, which may indicate that they received too much assistance during this initial administration. The issue is being explored through interviews with participating teachers, but no clear explanation has yet been identified. It is, however, expected that this type of uncertainty will be reduced when the test is administered to a larger and more controlled sample in the main study. Particularly if, in a future study, an additional test administration is included at an earlier stage to allow teachers and students to become familiar with the test material.

7.1.1 Feedback for teachers

The criterion-based design of test material has also been used to provide teachers with feedback at student level. In addition to students' scores on the subscales and the expected total score for the overall *MatematikProfilen* (MP) material, the feedback shows which categories the students are placed in at each of the three measurement points. An anonymised example from one of the project's grade 1 classes is presented in Tabel 3. The test results have been rounded to whole numbers to enhance usability for teachers. Here, it is clearly evident that a substantial change has occurred from Part B to Part C in the number of students placed in the two categories *Beyond expectations* and *Far beyond expectations* (green and blue highlighting).

Tabel 3. Example of feedback for participating teachers at student level

1A		1B		1C	
Total score	Converted score	Total score	Converted score	Total score	Converted score
12	36			9	36
12	36			12	49
13	39	10	25	13	53
13	39	11	28	13	53
13	39	12	31	14	55
13	39	13	34	14	55
14	42	13	34	14	55
14	42	14	38	14	55
15	46	15	42	15	58
15	46	15	42	16	59
15	46	15	42	16	59
15	46	15	42	16	59
16	49	15	42	16	59
16	49	16	46	17	61
16	49	16	46	17	61
17	53	16	46	17	61
17	53	17	50	17	61
18	56	17	50	18	62
18	56	17	50	18	62
18	56	18	54	18	62
19	59	18	54	19	64
20	62	18	54	19	64
20	62	19	57	19	64
21	64	19	57	20	65
21	64	22	66	21	66

In Table 3, the teacher of the class can, for example, see that two students in subtest 1B achieved a converted score on the overall MP scale of 34 points. This corresponds to the category *In progress*. In subtest 1C, the first student has moved from this category to *Beyond expectations*, with a score corresponding to 53 on the overall MP scale, while the second student is now in the category *Far beyond expectations*, with a score corresponding to 59 points on the overall MP scale. For the criterion-based feedback available to teachers, this corresponds to a description of the students' level of mathematical attainment in number and algebra:

Subtest 1B	Students in this category can typically add quantities up to 20 in everyday contexts, write numbers up to 20 in order of magnitude, and continue simple numerical or figural patterns. They are typically not yet able to explain strategies for addition and subtraction, and they still have a limited understanding of, and proficiency in using, the base-ten number system. (Gyldendal, our translation)
Subtest 1C	Students in this category can typically write numbers up to 100 in order of magnitude, add quantities up to 100 in everyday contexts, and continue simple numerical and figural patterns. They typically have a nuanced understanding of the base-ten number system and a good understanding of strategies for addition and subtraction. (Gyldendal, our translation)

7.1.2 Design-related challenges associated with test administration

The rolling implementation meant that compromises had to be made in the use of the test material. Students who began in the second half of grade 5 were given the grade 5 test even after progressing to grade 6 following the summer break. Since test material is designed for use at the end of the school year, this could potentially lead to ceiling effects. Our data show no evidence of ceiling effects; however, with a larger sample we cannot rule out that this could become an issue. We are therefore currently in the process of further developing the test material so that it can be used across grade levels. This is done by testing students in grade 1–5, with all classes completing one subtest corresponding to their current grade level and one subtest corresponding to the subsequent grade level within the same week. The assumption is that, within this one-week period, students will not have learned enough for any change to be measurable in test scores. In this way, it becomes possible to construct a conversion model that allows us to compare student scores across grade levels with the MP test material. A test material that allows for such cross-grade comparisons does not currently exist.

Although we will soon have a test material that allows for testing across grade levels, the plan is that all participating teachers in a forthcoming SeSam project will begin at the start of the school year, either in September or October, in order to avoid a summer break falling with either the control or the intervention period.

7.2 Teachers perceived outcomes based on three interviews

The teachers' statements point to a project that has left a lasting mark—both in terms of professional and subject-related development, collegial collaboration, and personal impact. Teacher A clearly experienced being challenged and has rediscovered a subject-related curiosity that also carries over into ambitions for students' learning. At the same time, encounters with colleagues' teaching and with their own video recordings have created a safe, professional and subject-focused learning space in which reflection is not about the individual, but about teaching as a shared object. Teacher B emphasises the rare, yet important, aspect of personally receiving subject-related input and engaging systematically in inquiry. At the same time, there is an attentiveness to students' experience of the teaching—and a perceived need to strike a balance between subject-focused depth and learning within familiar frameworks. Teacher C articulates the value of collaboration and an external professional perspective, while also expressing a critical attentiveness to whether the new approach actually improves learning for all students.

Across the interviews, it becomes clear that the NySeSam project has contributed to a deeper awareness of the structure, language, and purpose of teaching. The teachers highlight personal gains, collegial collaboration, and a changed attentiveness to students' learning as key outcomes of their participation. At the same time, the interviews point to the importance of creating sustained spaces for reflection. They also highlight the need for support to embed and consolidate these approaches in teachers' day-to-day routines if this form of teacher professional development is to continue beyond the conclusion of the project.

8 Discussion: What do we carry forward?

Through the implementation and subsequent analysis of the NySeSam pilot project, we have gained substantial insights that not only contribute to an understanding of the project's immediate effects but also provide a foundation for future development and scaling. These insights point toward a set of principles—practical, subject-related, and test-related alike—that are crucial if the project's elements are to be integrated meaningfully and sustainably into school practice.

A clear takeaway from the project is that development initiatives such as NySeSam require time. Teacher professional development is a process that unfolds over months, and similar initiatives should therefore be planned to run over longer periods of time. If the duration is too short, there is a risk that reflection and long-term consolidation will be limited. In line with this, it has proven meaningful to conduct at most one supervision cycle per month—consisting of lesson planning, teaching, and subsequent video-based supervision—as this allows time for preparation, experimentation, and properly reflecting on and thinking things through.

Collaboration in teams of teachers has likewise proven to offer particular opportunities. When several teachers participate together—especially when they teach the same grade level—a shared subject-specific frame of reference is established along with a collegial space for the exchange of experience that strengthens both a sense of safety and a culture of reflection. Workshops have functioned as a key moment in which participants have connected theory and practice, shared insights, and jointly developed a shared understanding of the ideas and concepts at the core of the project.

Three central didactic approaches have been particularly significant in NySeSam: video recording of teaching, workshops focused on sharing planning and reflection, and the use of structured problem-solving as a framework for inquiry-based teaching. The combination of these approaches has given teachers the opportunity to both experiment with new content and to develop a more inquiry-based, student-centred perspective on their own practice. The collaborative workshop format, in which teaching is made a shared object of analysis, has been particularly effective—and should be carried forward in future iterations.

The testing component of the project has provided valuable insights into students' learning progression, while at the same time raising important considerations about how assessments can be conducted in ways that are valid, meaningful, and sustainable. First, it has been a strength that the tests could be administered in under one hour, thereby minimising the burden on both students and teachers. Second, experience shows that students should not be exposed to the same test repeatedly, as this may affect validity. Finally, it is crucial that teachers are given access to their own students' test results, so that these can be used actively in their teaching practice and potentially replace other local tests. This increases the usefulness of the assessment and reduces the risk that testing is experienced as an additional burden.

Taken together, the reflections above indicate that NySeSam has not merely functioned as a bounded development initiative, but as a practice-based learning environment that, going forward, can form the basis for broader and more systematic capacity building in mathematics education.

References

Artigue, M. (2014). *Perspectives on design research: The case of didactical engineering*. In A. Bikner-Ahsbahs, C. Knipping, & N. Presmeg (Eds.), *Approaches to qualitative research in mathematics education: Examples of methodology and methods* (pp. 467–496). Springer Netherlands. https://doi.org/10.1007/978-94-017-9181-6_17

Asami-Johansson, Y. (2021). *Conditions and constraints for transferring Japanese structured problem solving to Swedish mathematics classroom*. *Recherche en Didactique des Mathématiques, 41*(3), 347–391.

Barquero, B., & Bosch, M. (2015). *Didactic engineering as a research methodology: From fundamental situations to study and research paths*. In A. Watson & M. Ohtani (Eds.), *Task design in mathematics education: An ICMI study 22* (pp. 249–272). Springer. https://doi.org/10.1007/978-3-319-09629-2_8

Barth-Cohen, L. A., Little, A. J., & Abrahamson, D. (2018). Building reflective practices in a pre-service math and science teacher education course that focuses on qualitative video analysis. *Journal of Science Teacher Education, 29*(1), 83–101. https://doi.org/10.1080/1046560X.2018.1423837

Eldridge, S. M., Chan, C. L., Campbell, M. J., Bond, C. M., Hopewell, S., Thabane, L., & Lancaster, G. A. (2016). CONSORT 2010 statement: Extension to randomised pilot and feasibility trials. *BMJ, 355*, i5239. https://doi.org/10.1136/bmj.i5239

Fox, A., & Poultney, V. (2020). Teacher professional learning through lesson study: Teachers' reflections. *International Journal for Lesson and Learning Studies, 9*(4), 397–412. https://doi.org/10.1108/IJLLS-03-2020-0011

Fullan, M., & Hargreaves, A. (2016). *Bringing the profession back in: A call to action*. Learning Forward.

Fullan, M., & Hargreaves, A. (2012). Reviving teaching with "professional capital." *Education Week, 31*(33), 30, 36. https://www.edweek.org/policy-politics/opinion-reviving-teaching-with-professional-capital/2012/06

Gee, D., & Whaley, J. (2016). Learning together: Practice-centred professional development to enhance mathematics instruction. *Mathematics Teacher Education and Development, 18*(1), 31–52.

Gelfuso, A., & Dennis, D. V. (2017). Video as text of teaching: Toward more deliberate literacy field experience supervision. *The Teacher Educator, 52*(1), 57–74. https://doi.org/10.1080/08878730.2016.1247203

Groves, S., Doig, B., Vale, C., & Widjaja, W. (2016). Critical factors in the adaptation and implementation of Japanese lesson study in the Australian context. *ZDM Mathematics Education, 48*(4), 501–512. https://doi.org/10.1007/s11858-016-0786-8

Koh, J. H. L. (2015). The use of video technology in pre-service teacher education and in-service teacher professional development. In J. M. Spector (Ed.), *The SAGE encyclopedia of educational technology* (pp. 142–145). SAGE Publications.

OECD (2023). *PISA 2022 assessment and analytical framework*. PISA. OECD Publishing. https://doi.org/10.1787/dfe0bf9c-en

Rasmussen, K. & Færch, J. V. (2021). Ny SeSam i Matematik – Design og afprøvning af et eksperiment med videostøttet sparring af matematiklæreres vedvarende professionelle udvikling (ansøgning til TrygFonden) [New SeSam in Mathematics: Designing and Piloting a Video-Supported Coaching Intervention for Mathematics Teachers' Sustained Professional Development (Application to TrygFonden)]

Rasmussen, K., & Isoda, M. (2019). The intangible task: A revelatory case of teaching mathematical thinking in Japanese elementary schools. *Research in Mathematics Education, 21*(1), 43–59. https://doi.org/10.1080/14794802.2018.1555714

Rogers, P. J. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation, 14*(1), 29–48. https://doi.org/10.1177/1356389007084674

Rüschenpöhler, L. (2024). Effectiveness and equity in primary science and mathematics education in the Nordic countries: A review of the book by Teig et al. (2024). *Nordic Studies in Education, 44*(3), 245–249. https://doi.org/10.23865/nse.v44.6691

Souma, K. (2017). Sugaku Mondaikaiketsu no Guidebook [The guidebook for the lessons with problem solving approach]. Tokyo: Meijitosho.

Sowder, J. T. (2007). *The mathematical education and development of teachers*. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (pp. 157–223). Information Age Publishing.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. Free Press.

Tarantini, E. (2021). *360°-video reflection in teacher education: A case study*. In D. G. Sampson, D. Ifenthaler, & P. Isaías (Eds.), *Proceedings of the 18th International Conference on Cognition and Exploratory Learning in the Digital Age (CELDA 2021)* (Virtual, October 13–15, 2021). International Association for Development of the Information Society. https://files.eric.ed.gov/fulltext/ED621339.pdf

Tülüce, H. S., & Çeçen, S. (2018). The use of video in microteaching: Affordances and constraints. *ELT Journal, 72*(1), 73–82. https://doi.org/10.1093/elt/ccx028

Walkoe, J., Sherin, M. G., & Elby, A. (2020). Video tagging as a window into teacher noticing. *Journal of Mathematics Teacher Education, 23*(4), 385–405.

Østergaard, C. H. & Færch, J. V. (2022). Ny SeSam i Matematik – Optimering af design (supplerende ansøgning til TrygFonden) [New SeSam in Mathematics: Design Optimization (Supplementary Application to the TrygFonden)]

9 Appendix

9.1 Appendix A: Literature review

Introduction

The practice of teacher reflection has long been recognised as a central pillar of teacher professional development and instructional quality (Schön, 1983). Within the field of mathematics education, reflective practices have gained increasing interest as part of collaborative models such as lesson study, where teachers engage in systematic cycles of planning, observing, and analysing students' learning.

In a lesson study cycle, teachers collaboratively delve into the content - of a mathematical subject area - focusing on a specific student learning objective. They then design and implement a lesson, which is followed by a phase of critical reflection. This process initiates subsequent cycles of studying, planning, enacting, and reflecting. Cobb et al. (2021) highlight the value of this method, noting "[l]esson study has the potential to support teachers' development of ambitious and equitable practices because it allows teachers to collaboratively investigate how to achieve a mathematical agenda by building on students' current thinking" (p. 120).

Lesson study is inherently student-centric, focusing primarily on student learning throughout its cycles. At every stage, activities are designed to focus teachers' attention on students' learning processes and how these are shaped by mathematical content and instructional strategies. Research shows that teachers who participate in collaborative lesson study communities develop a heightened awareness of student equity and inclusion, becoming more attentive to diverse learners and their needs (Aas, 2022, 2023; Aas et al., 2024; Schipper et al., 2020).

A critical component of the lesson study cycle is the research lesson—meticulously planned, executed, and analysed to understand and enhance student learning. Sometimes, these lessons are conducted without multiple iterations typical of full lesson study cycles; in such cases, we refer to them as open lessons. An open lesson involves presenting a comprehensive lesson plan, teaching it with students, and having observers carefully monitor specific aspects of student learning. Afterwards, discussions focus on analysing and reflecting on how effectively the lesson achieved its instructional goals and influenced student understanding of specific mathematical content.

Review I: Teacher reflections

The focus of this literature review is to explore how teacher reflections—individual, collective impact the quality of mathematics instruction. Specifically, this review synthesises findings from 11 peer-reviewed studies, highlighting key mechanisms through which reflection fosters changes in teachers' instructional practices, mathematical thinking, and professional identities.

Teacher reflection is multifaceted, encompassing personal introspection, collegial discussion, engagement with student thinking, and the interpretation of classroom evidence (Rodgers, 2002). Reflection may be guided by various tools—such as video, blackboard notes, or structured observation rubrics—and shaped by cultural, institutional, and pedagogical contexts. The reviewed literature captures this diversity while consistently underscoring reflection as a vital driver for learning and transformation in mathematics teaching.

In nearly all studies examined, reflection is portrayed as an indispensable element of lesson study (LS), especially in the Japanese model, which has become influential globally (Fujii, 2016; Ding et al., 2024). According to Rasmussen (2016), post-lesson reflections offer a unique paradidactic infrastructure for teachers to analyse their practices. These reflections allow teachers to develop new knowledge linking mathematical content with pedagogical strategies. He argues that collaborative reflection supports professional growth by helping teachers critically examine their choices and focus on student learning.

Asami-Johansson (2021) examines open lessons in Japan, showing how reflection is integrated into teachers' professional development. She investigates how Japanese mathematics teachers collaboratively produce and share didactic knowledge through post-lesson reflection meetings, also known as open lessons. The crucial aspect of this study is the dialectic between specific and generic levels of focus in teachers' reflections on observed teaching practices. This includes specific didactic techniques for achieving mathematical goals and broader pedagogical issues, such as realising the objectives of mathematics education. As pointed out in Asami-Johansen (2021), teacher reflection is important because it allows teachers to develop and share theoretical knowledge about their teaching practice. In relation to this, Rasmussen (2016) found that different institutional preferences, such as prospective teachers, in-service teachers, and teacher educators, who participate in the post-lesson discussions, provide new insights for the participants. Asami-Johansen (2021) also concluded that positive aspects of teacher reflection include the sharing of knowledge beyond a particular theme or domain and enhancing the theoretical block of teacher knowledge. She also mentions that the reflection sessions exemplify how Japanese teachers connect generic theories explicitly to their teaching practice. This contributes to developing shared knowledge about teaching through professional scholarship.

Similarly, Ding et al. (2024), in their synthesis of more than 150 Lesson Study studies, found that teachers' reflections often shift from focusing on teaching actions to analysing students' learning. They argue that productive reflection should centre on students' mathematical thinking and learning rather than on teachers' instructional performance. They further highlight that, particularly in Japan, knowledgeable others consistently direct post-lesson discussions towards children's ideas and mathematical work—a practice that has proven more difficult to establish in Lesson Study implementations outside Japan. Conversely, they found that teachers in the U.S. often reflected on social aspects of the LS process rather than analysing student learning. Bakker et al. (2022) and Bjuland & Mosvold (2015) found similar patterns in the Netherlands and Norway, respectively, where teachers did not observe students' thinking and learning during the research lesson.
Despite these challenges, Ding et al. (2024) highlight several positive impacts of effective reflection. For instance, Vermunt et al. (2019) found that reflections helped London teachers deepen their mathematical and pedagogical knowledge, particularly among less experienced teachers. However, the success of reflection also hinges on the collaborative culture of the LS team. As noted by Lee and Tan (2020), a supportive environment that encourages collegial critique enhances the

quality of reflection and teacher learning. In contrast, Sekao and Engelbrecht (2022) described how over-critical feedback during debriefing sessions led some teachers to feel attacked and disengaged from the process.

Review II: Video supervision

This review highlights the central role of teacher reflection in fostering instructional quality in mathematics education. Across all studies, structured and sustained reflection—especially within collaborative models such as lesson study—emerges as a key lever for pedagogical change. Video recordings and blackboard notes further serve to anchor discussions in concrete classroom evidence, supporting shifts in attention from teacher behaviour to student learning. Effective reflection is not automatic; it depends on professional structures, cultural norms, and facilitation tools. When thoughtfully designed and supported, reflection not only improves instructional strategies but also cultivates deeper professional identities and shared praxeologies among teachers. Future research should continue to explore how reflection can be sustained beyond short-term interventions and scaled across diverse educational contexts.

Video-based supervision has emerged as a powerful tool in mathematics teacher education, both in pre-service and in-service contexts. Rooted in the development of professional noticing, reflection, and pedagogical reasoning, video recordings allow educators to revisit and analyse classroom interactions with a depth not easily achieved through memory alone. This review synthesises empirical findings from 21 peer-reviewed studies and is structured around five thematic strands: (1) professional vision and noticing, (2) structured reflection and guided analysis, (3) digital and online environments, (4) implementation challenges, and (5) key principles for instructional design.

Video recordings of classroom teaching serve as a powerful mediating tool that enhances teacher reflection by allowing practitioners to observe their own instructional practices from a more objective distance. Several studies in the review demonstrate how video enables teachers to revisit classroom events, pause, analyse, and discuss with colleagues—activities that support metacognitive awareness and pedagogical shifts (Tarantini, 2021; Fox & Poultney, 2020; Gee & Whaley, 2016). For example, Tarantini (2021) explores the use of 360-degree videos in teacher education, highlighting their potential to provide multiple perspectives on teaching performance. Tarantini (2021) refers to studies showing that the immersive experience allows pre-service teachers to re-experience their teaching from different angles, enabling a more comprehensive analysis of teacher-student interactions. According to these studies, as cited in Tarantini (2021), this can lead to an improved understanding of how to conduct constructive and fruitful interactions with students. Furthermore, video-based reflection, especially when combined with social video learning (SVL), can facilitate situated learning, i.e. social collaborative knowledge construction, according to Gaudin & Chaliès (2015), as cited in Tarantini (2021). Similarly, in the study by Fox and Poultney (2020), teachers valued the opportunity to record and review their lessons as part of the lesson study process. Video was used to revisit classroom interactions, allowing teachers to analyse student responses and thinking. These recordings enabled repeated analysis and supported focused, collaborative discussions about student learning. While video was not the central tool for reflection, it served as a means for stepping back from the teacher role and engaging with practice. The non-judgmental setting of lesson study created space for this kind of inquiry and professional dialogue.

Gee and Whaley (2016) found that video-based reflection supported teachers in analysing their own instructional strategies, particularly questioning techniques and student thinking. Teachers selected short clips from their lessons, reflected individually, and then discussed them in peer groups. This process led to greater awareness of student responses and encouraged deeper, collaborative inquiry into teaching practices. Many teachers reported changes in their classroom instruction, including increased use of higher-order questions and stronger focus on student discourse. The video discussions also fostered mutual support and professional dialogue, making video a key tool for promoting reflective, student-centred teaching.

Reflection practices are deeply shaped by cultural norms and professional traditions. Studies from Japan and cross-cultural collaborations provide valuable insights into how cultural values influence the content and structure of teacher reflection.

Shimizu and Kang (2022) and Asami-Johansson (2021) describe how Japanese lesson study is organised around carefully facilitated post-lesson discussions that draw on detailed student work analysis and blackboard notes. These reflections are not ad hoc but follow rigorous routines and cultural expectations for deep engagement. The study by Shimizu and Kang (2022) further shows that reflection is not merely an individual activity but is shaped by a school culture that values collaborative inquiry, structured planning, and long-term professional growth. Over time, teachers' reflections shift from general pedagogical concerns to more nuanced discussions of students' thinking and mathematical understanding.

Schwarts and Karsenty (2020) investigate and analyse how Israeli teachers reflect on their own practice after watching a Japanese open lesson video. Cultural differences act as a catalyst for this reflection: Japanese practices—such as patience, structured pacing, and a focus on a single task—stand in clear contrast to what the authors describe as a hectic classroom atmosphere in Israel, where teachers often feel pressured to move quickly and cover a wide range of material to prepare students for high-stakes exams. This contrast opens space for what the authors refer to as vicarious reflection, where observing unfamiliar practices leads teachers to reconsider their own values and methods.

The reviewed studies also emphasise variation in the depth of teacher reflection. Ono and collegues (2013) developed a coding scheme to assess the depth of teachers' reflections during post-lesson discussions. By categorising reflections as descriptive, analytical, or interpretive, the scheme allows researchers and educators to identify the current level of reflection and track its development over time. The finding that reflections tend to begin descriptively aligns with the idea that teachers initially focus on concrete observations and experiences. However, the study also highlights the potential for structured facilitation to guide teachers toward more analytical and interpretive insights.

Richit and Tomkelski (2022) present a nuanced picture of how reflection during lesson study cycles enables teachers to reconceptualise teaching as an intentional educational act, a form of social commitment, and a process of pedagogical transformation. Some teachers expressed a shift in perspective, such as a teacher who emphasised a move toward more student-centred teaching. Others focused on the value of structured planning and evaluation. The reflections ranged from concrete instructional strategies to broader understandings of teaching's social and ethical purpose. The authors also note that, for many participants, the first experience with lesson study required significant adjustment before deeper reflection could occur, with reflection within lesson study being shaped by individual factors and specific contexts.

Tyskerud, Bjuland, and Mosvold (2023) further show how "commognitive conflicts" in discussions challenge teachers to refine their language and conceptual tools, thus deepening their pedagogical discourse. These conflicts help teachers reflect on their assumptions and shift from routine teaching to more explorative methods. By identifying these conflicts, teachers can work together to improve their teaching. The study highlights the importance of a knowledgeable other who can help teachers recognise and address these conflicts, leading to long-term improvements in their teaching practice.

Central to many studies is the potential of video to cultivate professional vision—the ability to notice and interpret significant moments in teaching and learning. Estapa and Amador (2023), in a qualitative metasynthesis of 49 studies, found that video prompts were particularly effective in developing teachers' awareness of student thinking. They caution, however, that "video does not produce learning" in itself but serves as a medium that "supports learning toward specific goals" (p. 106). This aligns with Star and Strickland's (2008) experimental study, which demonstrated that pre-service teachers exposed to video-based materials significantly enhanced their ability to recognise mathematically important classroom events. These findings challenge the notion that noticing is innate, instead supporting the idea that "teachers can learn to notice with practice" (Star & Strickland, 2008, p. 124). Complementing this view, Walkoe et al. (2020) introduced video tagging as a means of training teachers to isolate and reflect on pivotal pedagogical moments. Through tagging, pre-service teachers moved beyond mere description to engage in deeper analytical practices, thereby strengthening their ability to recognise and interpret students' mathematical thinking.

Structured reflection, particularly when guided by supervisors or peer dialogue, emerges as a cornerstone of successful video integration. Karakaş and Yükselir (2021), working with pre-service language teachers, showed that video-mediated microteaching combined with structured discussions allowed participants to identify aspects of their teaching that would otherwise go unnoticed. Video, they argue, "provided deeper insights into teaching practices" (p. 167), facilitating the development of pedagogical awareness. Recursive video-stimulated recall is another method with strong empirical support. Lutovac et al. (2015) demonstrated how repeated viewings fostered metacognitive growth, enabling one pre-service teacher to articulate increasingly nuanced insights into her teaching. Similarly, Barth-Cohen et al. (2018) found that structured, qualitative analysis of video encouraged pre-service teachers to think more like researchers—identifying patterns, forming interpretations, and revising their understandings over time. The role of the facilitator is especially critical. Gelfuso and Dennis (2017) describe video as a "text of teaching," a shared object for interrogation that, when scaffolded properly, shifts post-lesson dialogue from vague impressions to precise, evidence-based conversation. Yet without such scaffolding, video can fail to achieve its full potential. As Tülüce and Çeçen (2018) observed, pre-service teachers without guidance tended to focus on superficial aspects, such as body language or voice projection, rather than instructional decisions. This is echoed by Osmanoğlu (2016), who found that only with structured support could teachers move "from noticing to interpreting" (p. 70).

With the expansion of remote learning, video has become even more integral to teacher education. Beilstein et al. (2021) found that in online learning communities, teachers interacted with video content in complex, often invisible ways, highlighting the need for multi-modal engagement and purposeful integration into digital platforms. Güler and Çelik (2022) reported that e-mentoring using video clips improved novice mathematics teachers' ability to analyse classroom discourse and student thinking. Participants became more attuned to subtle indicators of learning, and "the video-supported e-mentoring process enhanced their awareness of students' thinking" (p. 88). Likewise, Svensson (2023) showed that pre-service teachers working with early childhood classroom recordings improved their ability to distinguish between every day and mathematical activity, an essential skill for foundational mathematics teaching.

Despite the many documented benefits, several challenges complicate the use of video in teacher education. Emotional discomfort is a recurrent theme. Participants in Tülüce and Çeçen's (2018) study reported anxiety about being filmed, which hindered their authenticity and self-reflection. Koh (2015) identifies a broader structural issue: video-based programs often lack theoretical coherence or clear learning objectives, reducing their effectiveness. Furthermore, the process of analysing video is time-consuming and cognitively demanding. As Barth-Cohen et al. (2018) point out, without sufficient time and support, the analytical work required of teachers can become overwhelming. Gelfuso and Dennis (2017) similarly warn that poorly facilitated discussions may drift toward vague or overly critical commentary.

The literature converges on several design principles for effective use of video in mathematics teacher education. First, facilitation is paramount. Whether through structured discussion prompts, video tagging, or recursive viewing, scaffolding supports deeper engagement and mitigates surface-level reflection. Second, the content of the video must align with clear learning objectives. Seago et al. (2018) advocate for a "video in the middle" design, where video serves as the central artifact around which professional development tasks are structured. Purposeful sequencing of pre- and post-viewing activities, they argue, is essential to achieving meaningful outcomes. Third, collaboration is key. Video is most effective when used in a collective setting where multiple perspectives can challenge and enrich interpretation. This collective analysis strengthens teachers' pedagogical reasoning and supports professional learning as a social endeavour.

While the previous sections have highlighted the significance of facilitation and structured guidance in video-based supervision, the central role of the supervisor, often conceptualised as the knowledgeable other, warrants deeper examination. Supervisors do not merely support reflection; they shape the interpretive frameworks through which pre-service and in-service teachers make sense of classroom interactions. In the following section, we delve more closely into the characteristics, functions, and challenges of the knowledgeable other, drawing on recent empirical studies that illuminate how this role influences both teacher learning and the quality of collaborative inquiry. This focused exploration allows us to unpack the pedagogical, relational, and epistemological dimensions of supervision within mathematics education.

Review III: The role of the knowledgeable other in teacher professional development

The role of the knowledgeable other has emerged as a central component in teacher professional development, particularly within the framework of lesson study in mathematics education. Whether referred to as mentors, facilitators, or outside experts, knowledgeable others offer critical support to teacher learning by contributing pedagogical, theoretical, and subject-specific insights. This review synthesises empirical findings on the nature, influence, and complexities of the knowledgeable other role, with particular attention to how these individuals support collaborative inquiry, mediate between theory and practice, and foster pedagogical content knowledge.

The functions of the knowledgeable other are rich and varied. Takahashi (2014) identifies three key responsibilities: introducing new research-based perspectives, linking theory to classroom practice, and scaffolding reflective dialogue. Similarly, Watanabe (2005) underscores that knowledgeable others enrich the process by "provid[ing] a different perspective when reacting to the lesson study work of the group" and by bringing insights from other lesson study contexts (p. 1). Rather than prescribing solutions, effective the knowledgeable others stimulate teacher reflection through well-posed questions and guided discussion. As Groves and Doig (2016) note, the strength of a knowledgeable other lies in their capacity to prompt deep thinking without directing outcomes. Murata (2011) further adds that the knowledgeable other brings an "external authority" that helps legitimise professional discussions and keep them anchored in goals related to student learning and curricular coherence. Beyond facilitating technical adjustments, the knowledgeable others also shape teachers' beliefs. According to Seino and Foster (2021), final reflective comments from the knowledgeable others in Japanese lesson study cycles often prompted significant interpretive shifts among participating teachers. These interventions offered not just feedback but moments of reframing, allowing teachers to reconsider assumptions about mathematics teaching and learning. Importantly, Abtahi and collegues (2017) propose a reconceptualisation of the knowledgeable other through the lens of a "multi-directional Zone of Proximal Development." Here, the knowledgeable other is not a top-down figure but a co-learner, engaging in dialogic meaning-making alongside teachers. As they argue, "the more knowledgeable other is positioned as a co-learner rather than a directive figure" (p. 8), challenging traditional hierarchies in professional learning.

Another prominent function of the knowledgeable other lies in supporting teachers' professional noticing—the ability to attend to and interpret classroom interactions in ways that inform instruction. Amador and Weiland (2015) found that while pre-service teachers primarily focused on surface-level observations, the knowledgeable others brought attention to broader aspects such as task structure, pacing, and alignment with mathematical goals. This broadened lens helps guide novices toward more nuanced pedagogical understanding. The power of the knowledgeable other to structure and elevate reflection is also evident in Seino and Foster's (2021) categorisation of the knowledgeable other commentary into seven thematic areas, from addressing the didactical value of content to facilitating teacher growth. They assert that "the quality of these final comments is critical to the learning of the LS participants" (p. 507), suggesting that the knowledgeable others shape not only what teachers notice, but how they learn from it. Echoing this, Takahashi and McDougal (2016) emphasise that effective the knowledgeable others model an inquiry stance. By prompting teachers to investigate student thinking and revise instructional strategies, they help ensure that lesson study remains rooted in evidence-based teaching improvement.

One of the most cited benefits of the knowledgeable other is their ability to bridge the often-discussed gap between educational theory and teaching practice. Murata (2011) illustrates how the knowledgeable others translate research, standards, and developmental theories into the lived realities of the classroom. Without this mediation, lesson study risks becoming either abstract or merely procedural. In adapting Japanese models to the U.S., Takahashi and McDougal (2016) coined the term Collaborative Lesson Research, emphasising structured implementation and expert facilitation. Their experience shows that the presence of a knowledgeable other can transform collaborative planning into a rigorous professional learning process. The knowledgeable others also influence planning. In their study of lesson design, Bilge and Dede (2024) report that knowledgeable others helped pre-service teachers anticipate student responses and fine-tune pedagogical strategies: "The researcher's contributions helped PSTs better understand students' likely responses and adapt their lesson accordingly" (p. 7). As Groves and Doig (2016) affirm, such contributions deepen the reflective process by helping teachers "make links between pedagogical decisions and student learning outcomes" (p. 54), ensuring that reflection moves beyond the descriptive and into the analytical.

Despite the documented benefits, the role of the knowledgeable other is not without tensions. Fujii (2016) cautions that over-reliance on external experts may limit teachers' ownership of the inquiry process. Similarly, Bocala (2015) warns that expert input, while helpful, must be balanced to avoid undermining teacher agency. The cultural context in which lesson study is embedded also plays a critical role. In Japan, the knowledgeable others often hold senior teaching positions and bring both content expertise and institutional legitimacy. In other countries, the role is sometimes assumed by researchers or facilitators, who must also navigate administrative and logistical responsibilities (Watanabe, 2005). This shift may blur boundaries between collegial inquiry and external evaluation. To clarify expectations, Watanabe (2005) outlines six traits of effective knowledgeable others, including deep content knowledge, experience with lesson study, and an openness to learning. He emphasises that a "learning stance"—a willingness to learn from teachers and students alike—is perhaps the most vital trait: "Knowledgeable others can learn from teacher practitioners of lesson study (and students)" (p. 3). A further challenge lies in preparation. As Takahashi (2014) observes, few formal programs exist to train the knowledgeable others; most develop their expertise through informal, experiential learning. This raises concerns about scalability and fidelity when lesson study is exported to new contexts. Ensuring that the knowledgeable others are adequately prepared remains a pressing need for international adaptations of lesson study.

The literature clearly demonstrates that the knowledgeable other plays a pivotal role in deepening the quality of teacher professional development. When enacted thoughtfully, the knowledgeable other serves as a bridge between theoretical frameworks and practical decisions, helping teachers develop both pedagogical and mathematical expertise. By guiding reflection, enriching planning, and challenging assumptions, the knowledgeable others contribute to lasting instructional improvement. However, their impact is mediated by context, relationship dynamics, and the presence—or absence—of appropriate training structures. Future research should aim to articulate clearer models for the preparation and support of knowledgeable others, particularly in non-Japanese settings. Moreover, reimagining the knowledgeable other as a co-learner and dialogic partner rather than an expert authority may foster more sustainable and egalitarian cultures of professional learning.

References (reviews)

Aas, H. K. (2022). Teachers talk on student needs: Exploring how teacher beliefs challenge inclusive education in a Norwegian context. *International Journal of Inclusive Education, 26*(5), 495–509. https://doi.org/10.1080/13603116.2019.1698065

Aas, H. K. (2023). Professional development for inclusive and adaptive education: Lesson study in a Norwegian context. *Professional Development in Education, 49*(3), 491–505. https://doi.org/10.1080/19415257.2020.1850509

Aas, H. K., Uthus, M., & Løhre, A. (2024). Inclusive education for students with challenging behaviour: Development of teachers' beliefs and ideas for adaptations through Lesson Study. *European Journal of Special Needs Education, 39*(1), 64–78. https://doi.org/10.1080/08856257.2023.2191107

Abtahi, Y., Graven, M., & Lerman, S. (2017). Conceptualising the more knowledgeable other within a multi-directional ZPD. *Educational Studies in Mathematics, 96*(3), 275–287. https://doi.org/10.1007/s10649-017-9768-1

Amador, J. M., & Weiland, I. S. (2015). What preservice teachers and knowledgeable others professionally notice during lesson study. *The Teacher Educator, 50*(2), 109–126. https://doi.org/10.1080/08878730.2015.1009221

Asami-Johansson, Y. (2021). The didactic notion of "mathematical activity" in Japanese teachers' professional scholarship: A case study of an open lesson. *REDIMAT – Journal of Research in Mathematics Education, 10*(1), 88–112. https://doi.org/10.17583/redimat.2021.4598

Bakker, C., de Glopper, K., & de Vries, S. (2022). Noticing as reasoning in Lesson Study teams in initial teacher education. *Teaching and Teacher Education, 113*, Article 103656. https://doi.org/10.1016/j.tate.2022.103656

Barth-Cohen, L. A., Little, A. J., & Abrahamson, D. (2018). Building reflective practices in a pre-service math and science teacher education course that focuses on qualitative video analysis. *Journal of Science Teacher Education, 29*(1), 83–101. https://doi.org/10.1080/1046560X.2018.1423837

Beilstein, S. O., Henricks, G. M., Jay, V., Perry, M., Bates, M. S., Moran, C. G., & Cimpian, J. R. (2021). Teacher voices from an online elementary mathematics community: Examining perceptions of professional learning. *Journal of Mathematics Teacher Education, 24*(3), 283–308. https://doi.org/10.1007/s10857-020-09459-z

Bilge, O., & Dede, Y. (2024). Researcher as knowledgeable other in the planning phase of a lesson study in mathematics. *International Journal of Mathematical Education in Science and Technology*. Advance online publication. https://doi.org/10.1080/0020739X.2023.2170288

Bjuland, R., & Mosvold, R. (2015). Lesson study in teacher education: Learning from a challenging case. *Teaching and Teacher Education, 52*, 83–90. https://doi.org/10.1016/j.tate.2015.09.005

Bocala, C. (2015). From experience to expertise: The development of teachers' learning in lesson study. *Journal of Teacher Education, 66*(4), 349–362. https://doi.org/10.1177/0022487115592032

Cobb, P., Jackson, K., Henrick, E., & Smith, T. M. (2021). *Systems for instructional improvement: Creating coherence from the classroom to the district office*. Harvard Education Press.

Ding, M., Huang, R., Pressimone Beckowski, C., Li, X., & Li, Y. (2024). A review of lesson study in mathematics education from 2015 to 2022: Implementation and impact. *ZDM–Mathematics Education, 56*(1), 87–99. https://doi.org/10.1007/s11858-023-01538-8

Estapa, A., & Amador, J. M. (2023). A qualitative metasynthesis of video-based prompts and noticing in mathematics education. *Mathematics Education Research Journal, 35*(2), 105–131. https://doi.org/10.1007/s13394-021-00378-7

Fox, A., & Poultney, V. (2020). Teacher professional learning through lesson study: Teachers' reflections. *International Journal for Lesson and Learning Studies, 9*(4), 397–412. https://doi.org/10.1108/IJLLS-03-2020-0011

Fujii, T. (2016). Designing and adapting tasks in lesson planning: A critical process of Lesson Study. *ZDM Mathematics Education, 48*(4), 411–423. https://doi.org/10.1007/s11858-016-0770-3

Gee, D., & Whaley, J. (2016). Learning together: Practice-centred professional development to enhance mathematics instruction. *Mathematics Teacher Education and Development, 18*(1), 31–52.

Gelfuso, A., & Dennis, D. V. (2017). Video as text of teaching: Toward more deliberate literacy field experience supervision. *The Teacher Educator, 52*(1), 57–74. https://doi.org/10.1080/08878730.2016.1247203

Chaliès, S., Gaudin, C., & Tribet, H. (2015). Exploiter la vidéo dans les dispositifs de formation des enseignants novices: Conceptualisation et discussion théoriques à partir d'une étude de cas en EPS. *Revue française de pédagogie, 193*, 5–24. https://doi.org/10.4000/rfp.4880

Groves, S., & Doig, B. (2016). *The role of the knowledgeable other in post-lesson discussions in lesson study*. In *Proceedings of the 40th Conference of the International Group for the Psychology of Mathematics Education* (pp. 315–322). International Group for the Psychology of Mathematics Education.

Güler, M., & Çelik, D. (2022). *Supporting novice mathematics teachers: The impact of e-mentoring on lesson analysis skills*. *Teaching and Teacher Education, 113*, Article 103658. https://doi.org/10.1016/j.tate.2022.103658

Karakaş, A., & Yükselir, C. (2021). Engaging pre-service EFL teachers in reflection through video-mediated team micro-teaching and guided discussions. *Reflective Practice, 22*(2), 159–172. https://doi.org/10.1080/14623943.2020.1860927

Koh, J. H. L. (2015). The use of video technology in pre-service teacher education and in-service teacher professional development. In J. M. Spector (Ed.), *The SAGE encyclopedia of educational technology* (pp. 142–145). SAGE Publications.

Lee, L. H. J., & Tan, S. C. (2020). Teacher learning in Lesson Study: Affordances, disturbances, contradictions, and implications. *Teaching and Teacher Education, 89*, Article 102986. https://doi.org/10.1016/j.tate.2019.102986

Lutovac, S., Kaasila, R., & Juuso, H. (2015). Video-stimulated recall as a facilitator of a pre-service teacher's reflection on teaching and post-teaching supervision discussion: A case study from Finland. *Journal of Education and Learning, 4*(3), 14–24. https://doi.org/10.5539/jel.v4n3p14

Murata, A. (2011). Introduction: Conceptual overview of lesson study. In L. C. Hart, A. S. Alston, & A. Murata (Eds.), *Lesson study research and practice in mathematics education* (pp. 1–12). Springer. https://doi.org/10.1007/978-90-481-9941-9_1

Ono, Y., Chikamori, K., & Rogan, J. M. (2013). How reflective are lesson study discussion sessions? Developing an instrument to analyze collective reflection. *International Journal of Education, 5*(3), 52–67. https://doi.org/10.5296/ije.v5i3.3847

Osmanoğlu, A. (2016). Prospective teachers' teaching experience: Teacher learning through the use of video. *Educational Research, 58*(1), 39–55. https://doi.org/10.1080/00131881.2015.1117321

Rasmussen, K. (2016). Lesson study in prospective mathematics teacher education: Didactic and paradidactic technology in the post-lesson reflection. Journal of Mathematics Teacher Education, 19(4), 301–324. https://doi.org/10.1007/s10857-015-9299-6

Richit, A., & Tomkelski, M. L. (2022). Meanings of mathematics teaching forged through reflection in a lesson study. *Eurasia Journal of Mathematics, Science and Technology Education, 18*(9), em2151. https://doi.org/10.29333/ejmste/12325

Rodgers, C. R. (2002). Seeing student learning: Teacher change and the role of reflection. *Harvard Educational Review, 72*(2), 230–253.

Schipper, T. M., van der Lans, R. M., de Vries, S., Goei, S. L., & van Veen, K. (2020). Becoming a more adaptive teacher through collaborating in Lesson Study? Examining the influence of Lesson Study on teachers' adaptive teaching practices in mainstream secondary education. *Teaching and Teacher Education, 88*, Article 102961. https://doi.org/10.1016/j.tate.2019.102961

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.

Schwartz, G., & Karsenty, R. (2020). "Can this happen only in Japan?": Mathematics teachers reflect on a videotaped lesson in a cross-cultural context. *Journal of Mathematics Teacher Education, 23*(5), 527–554. https://doi.org/10.1007/s10857-019-09438-z

Seino, T., & Foster, C. (2021). Analysis of the final comments provided by a knowledgeable other in lesson study. *Journal of Mathematics Teacher Education, 24*, 507–528. https://doi.org/10.1007/s10857-020-09468-y

Sekao, D., & Engelbrecht, J. (2022). South African primary mathematics teachers' experiences and perspectives about lesson study. *International Journal of Science and Mathematics Education, 20*(7), 1431–1453. https://doi.org/10.1007/s10763-021-10214-w

Shimizu, Y., & Kang, H. (2022). Discussing students' thinking and perspectives for improving teaching: An analysis of teachers' reflection in post-lesson discussions in lesson study cycles. *ZDM–Mathematics Education, 54*(2), 419–431. https://doi.org/10.1007/s11858-022-01371-5

Star, J. R., & Strickland, S. K. (2008). Learning to observe: Using video to improve preservice mathematics teachers' ability to notice. *Journal of Mathematics Teacher Education, 11*(2), 107–125. https://doi.org/10.1007/s10857-007-9063-7

Svensson, C. (2023). Using video-recordings to enhance pre-service teachers' abilities to identify mathematical activities in pre-school. *Cogent Education, 10*(1), Article 2194778. https://doi.org/10.1080/2331186X.2023.2194778

Takahashi, A. (2014). The role of the knowledgeable other in lesson study: Examining the final comments of experienced lesson study practitioners. *Mathematics Teacher Education and Development, 16*(1), 4–21.

Takahashi, A., & McDougal, T. (2016). Collaborative lesson research: Maximizing the impact of lesson study. *ZDM Mathematics Education, 48*(4), 513–526. https://doi.org/10.1007/s11858-015-0752-x

Tarantini, E. (2021). *360°-video reflection in teacher education: A case study*. In D. G. Sampson, D. Ifenthaler, & P. Isaías (Eds.), *Proceedings of the 18th International Conference on Cognition and Exploratory Learning in the Digital Age (CELDA 2021)* (Virtual, October 13–15, 2021). International Association for Development of the Information Society. https://files.eric.ed.gov/fulltext/ED621339.pdf

Tülüce, H. S., & Çeçen, S. (2018). The use of video in microteaching: Affordances and constraints. *ELT Journal, 72*(1), 73–82. https://doi.org/10.1093/elt/ccx028

Tyskerud, A., Bjuland, R., & Mosvold, R. (2023). Commognitive conflicts in mathematics teachers' pedagogical discourse in lesson study. *Mathematics Teacher Education and Development, 25*(1), 81–92.

Vermunt, J. D., Vrikki, M., van Halem, N., Warwick, P., & Mercer, N. (2019). The impact of Lesson Study professional development on the quality of teacher learning. *Teaching and Teacher Education, 81*, 61–73. https://doi.org/10.1016/j.tate.2019.02.009

Walkoe, J., Sherin, M. G., & Elby, A. (2020). Video tagging as a window into teacher noticing. *Journal of Mathematics Teacher Education, 23*(4), 385–405.

Watanabe, T. (2005). *Knowledgeable others: What are your roles and how do you become more effective?* Handout presented at the NWREL Lesson Study Symposium, Portland, OR. https://educationnorthwest.org/sites/default/files/handout13-1.pdf

9.2 Appendix B: Overview of participating schools

NySeSam-skema

Skole	KP-makker	Lærer	Klasse	Workshop	Samtale lektionsplan	Samtale lektion/video	Interview
2	CHOE	D	3. klasse	2	4	4	Ja
2	CHOE	E	3. klasse	2	3 (+ en video)	4	Ja
2	CHOE	B	4. klasse	2	4	4	Ja
2	CHOE	C	4. klasse	1	4	4	Ja
2	PSTH	A	1. klasse	2	5	5	Ja
2	PSTH	F	6. klasse	1,5	4	4	Ja
2	PSTH	G	5. klasse	2	4	4	Nej
2	Monica	H					
3	INMO	I	2. klasse	2	4	4	Ja
3	INMO	J	1.klasse	2	3	2	Ja
3	PSAS	K	3. klasse	2	4	4	Ja
3	PSAS	L	3. klasse	2	4	4	Ja
3	PSAS	M	3. klasse	2	2	2	Nej
3	SIGM	N	5. klasse	2	4	4	Ja
R3	SIGM	O	5. klasse	2	4	4	Ja
1	PSAS	P	3. 4. og 5. klasse	2	3	3	Nej
1	PSAS	Q	2. og 3. klasse	2	3	3	Nej
1	PSTH	R	5. og 6. klasse	2	3	3	Nej
1	PSTH	S	3. klasse	2	2	2	Nej
1	SIGM	T	1. og 2. klasse	2	2	1	Nej
1	SIGM	U	6. klasse	2	2	2	Nej
1		V	vejleder				

9.3 Appendix C: Information and cadence

NySeSam kort information og kadence

Kære NySeSam deltager, her er information og program for forløbet

I SeSam er det meningen, at vi sammen skal blive klogere på undervisning. Vi interesserer os rigtig meget for, hvad og hvordan børn lærer.

Aktivitet	Beskrivelse	Tidspunkt	Varighed
Test I	Vi kommer ud på skolen og hjælper med at tage testen	April	1
Test II	Vi kommer ud på skolen og hjælper med at tage testen	August (en uge efter skolestart)	1
Workshop I	Indføring i lektionsstudietankegangen Den korte lektionsplan Åben lektion og refleksion Forventningsafklaring og mød din 'NySeSam koshi' Formalia, herunder videosystemet.	Slut august	4
I: Forberedelse af lektionsplan	Lærer(team) arbejder med udarbejdelse af lektionsplan	Slut august	2
I: Samtale om plan	Ekstra forberedelse og sende plan (45 min), samtale om plan (30 min), tilretning af plan (15min)	September	1,5
I: Samtale om video af lektion	Optage og sende video (15 min), gense video (45min), samtale om video (30 min)		1,5
II: Forberedelse af lektionsplan			2
II: Samtale om plan			1,5
II: Samtale om video af lektion			1,5
Workshop II	Afklaring af generelle udfordringer Åben lektion og refleksion Udarbejdelse af lektionsplaner og diskussion		4
III: Forberedelse af lektionsplan		Oktober	2
III: Samtale om plan			1,5
III: Samtale om video af lektion			1,5
IV: Forberedelse af lektionsplan			2
IV: Samtale om plan			1,5
IV: Samtale om video af lektion			1,5
Test III	Vi kommer ud på skolen og hjælper med at tage testen	November	1
Interview			2
Ekstra tid			2
Samlet tidsramme			35

9.4 Appendix D: Interview guide

Interviewguide: NySeSam 2024

Workshop:

- Beskriv første workshop, hvor du/I blev introduceret til NySeSams grundlæggende ideer, organisation og teknisk set up.
    - o Hvordan oplevede du den åbne lektion / mockup lektion / video-lektionen?
- Beskriv den anden workshop, hvor du/I arbejdede videre med NySeSam.
    - o Hvordan oplevede du at udveksle forskellige erfaringer med andre deltagere?

Fokus på delelementer i NySeSam:

- Lektionsplan
    - o Hvordan var det at lave en lektionsplan?
    - o Havde du mulighed for at drøfte den med kolleger?
    - o Hvad er din oplevelse af at drøfte den med NySeSam-makker?
      (Tænk gerne på konkrete planer, dvs. noget kan jo opleves udfordrende, andre givtige…)
- Gennemførelse
    - o Hvordan har din oplevelse været med at udføre lektionen?

      (Tænk gerne på konkrete lektioner, dvs. noget kan jo opleves udfordrende, andre givtige…)
- Læring
    - o Hvad husker du særligt, du fandt ud af om elevernes matematiklæring og den undervisning, som du afholdt?
- Ramme
    - o Har du brugt nogle af din oplevelser / ny indsigt i elevlæring i anden undervisning eller i samtale med kollega? Har I haft mulighed for at samarbejde i team?
- Fremtid
    - o Hvordan tænker du, at I som faggruppe kan bruge jeres erfaringer?
- Hvilke temaer / forhold, husker du umiddelbart, har været på dagsordenen for dig og din makkers samtale?
    - o Undren over elevers matematiklæring?
    - o Undren over, hvordan undervisningen fungerer?
    - o Matematikfagligt fokus?

Organisering:

- Hvordan oplevede du **video-sparring** (virtuel sparring)?
- Hvordan vil du beskrive dit **samarbejde** med din NySeSam-makker fra KP?
    - o Kan du give eksempler på, hvordan I har samarbejdet?
    - o Beskriv, hvad du har taget med dig fra projektet.
- Hvordan har din oplevelse med den **tekniske side** af NySeSam været?
    - o Beskriv gerne, hvordan det var i starten, og hvordan det blev hen mod slutningen.

9.5 Appendix E: Dialogue guide

Supervisionsvejledning til NySeSam

Afsættet for NySeSam supervisionen kommer fra det japanske begreb om ”kenkyuu juygjo”, der dækker over en omfattende kultur og tankegang for udvikling og dissemimination af lærerviden. Med til kulturen hører kendskabet til, eller bestræbelsen efter at kende, den viden om undervisningens gennemførelse og fagdidaktiske indsigt, som er mest hensigtsmæssig for elevens læring.

En del af lektionsstudiekulturens kerne drejer sig om at observere en eller flere lektioners undervisning hvorom der gøres ekstra ud af forarbejdet (planlægningen) og efterarbejdet (refleksion, eftertanke). For sådanne ”studie-lektioner” er der altid defineret et klart og tydeligt **undersøgelsesformål**. Ved nogle studie-lektioner deltager en person som er ekstern i forhold til den klasse og den skole hvori lektionen foregår. Denne person, på japansk kaldet ’koshi’, som betyder noget i retning af ”lærd ven” eller ”vidende anden”, bidrager særligt i efterarbejdet med en nysgerrig interesse for hvilken elev-læring der kan ses i den givne lektion, i relation til det givne undersøgelsesformål.

Som supervisor i NySeSam er du lærerens ’koshi’, og bidrager til undersøgelsesformålet gennem supervision på to sammenhængende delelementer: først på en kortfattet lektionsplan og siden på en videooptagelse af den gennemførte lektion.

Første delelement: Den korte lektionsplan.

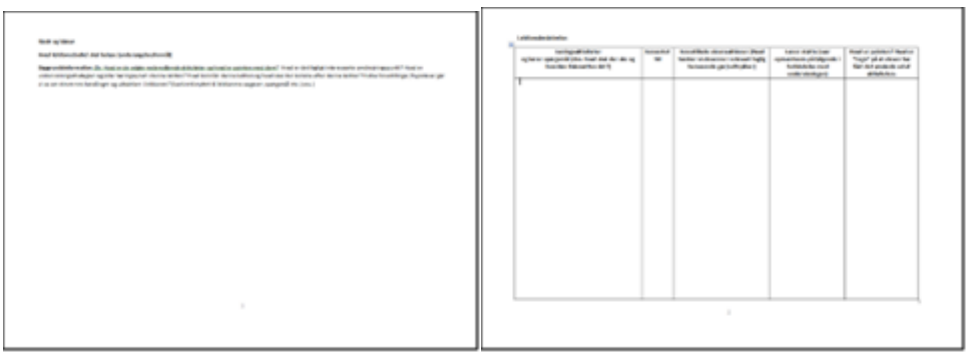
De vigtige processer er i forbindelse med nedskrivningen:

1) at supervisanten får tænkt over, og valgt noget som han/hun gerne vil blive klogere på i forhold til sin undervisning og elevernes læring. (Dette er undersøgelsesformålet)

2) at supervisanten ”sætter sig i elevernes sted”; overvejer og selv afprøver det som eleverne påtænkes at skulle gøre.

3) at supervissanten gør sig forestillinger om hvad eleverne kan tænkes at gøre/svare.

Disse planlægningsmæssige overvejelser nedskrives, evt. i stikordsform på to liggende A4-sider i følgende skabelon:



Anden delelement: Fælles refleksion over en videooptagelse af den gennemførte lektion.

Vær opmærksom på at der i NySeSam er udeladt mange facetter fra den omfattende lektionsstudie-kultur i almindelighed. NySeSam bruger nogle centrale elementer og tankegods fra LS, men påberåber sig i øvrigt ikke at give et dækkende billede af, hvad lektionsstudieri øvrigt kan siges at måtte omfatte.

Den implicitte normative dimension

Lektionsstudiekulturen har i Japan frembragt en erkendelse af, at det vigtigste i undervisningen er viden om elevers læring af et defineret indhold, der som er tæt relateret til et stofligt/fagligt område. Der er fokus på at eleverne kommer til orde i undervisningen, og at fællesskabet i klassen udnyttes til at få fremdraget så mange vinkler på det faglige som muligt. Læreren er således ikke en der fortæller/forklarer en hel masse, men overgiver til eleverne at undersøge, ræsonnere og præsentere deres tanker. Det vigtige er således elevens relation med stoffet, og læreren er mere den der orkestrerer elevernes/klassens arbejde som helhed, end en der interagerer med eleverne individuelt.

Lektionsstuder interesserer sig for elevers læring af noget, og vil søge at løse udfordringer i undervisningen gennem didaktiske og fagdidaktiske greb, i højere grad end gennem almene pædagogiske greb, såsom fx gruppesammensætninger, læringsstile mv.

Undervisningsdifferentiering kommer til udtryk gennem fremdragelse af forskellige elevers faglige perspektiver for klassefællesskabet, snarere end at udstikke individuelle opgaver på forskellige niveauer. Der søges opnået at den enkelte bidrager til fællesskabet, og der ønskes en kultur hvor også simple og fejlagtige forslag/ytringer anses som værdifulde skridt mod en mere fuldstændig/omfattende forståelse af den viden og kunnen der er ”på spil” i undervisningen.

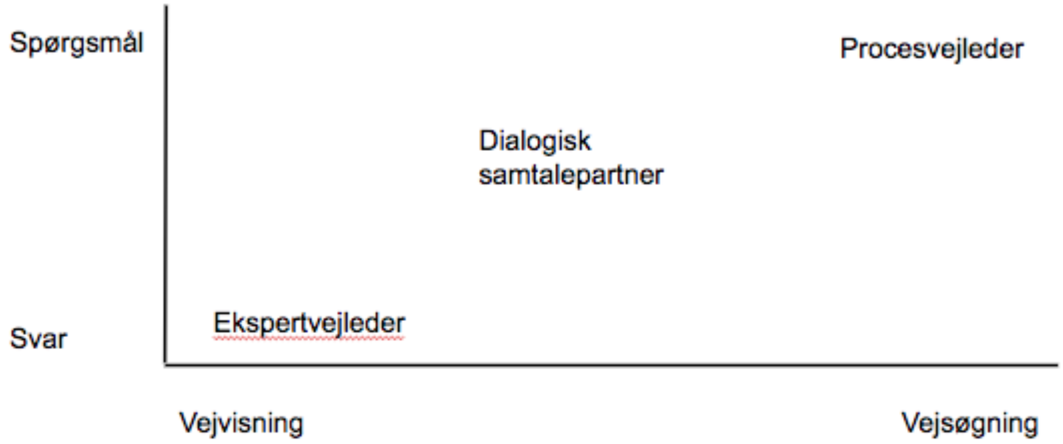
Motivations og engagementsmæssige udfordringer søges løst gennem et engagerende indhold, dvs. at undervisningen (og underviseren) faktisk har noget ”på hjerte”. At der er ”raison d'etre” for det eleverne skal lave. Megen undervisning i japanske lektionsstudier er derfor af en undersøgende og konkret karakter. Men det betyder ikke at al undervisning partout skal struktureres således.

Det vigtige er at undervisningen opbygges/indrettes således at eleverne kan lære det som er målet/meningen/pointen med undervisningen. Og at man som lærer/observatør kan komme til at se hvori elevernes læring består. Det er væsentligt at der tænkes over sammenhængen mellem pointen, læringsaktiviteterne og lærerens mulighed for at få indblik i læringen.

Supervision

Almindeligvis angives der to hovedretninger inden for supervisionsteorien. Den vejvisende og den vejsøgende.

De to retninger kan se illustreret herunder:



(Løw, 2009)

Når supervisor optræder som vejviser, vil han/hun som udgangspunkt udnytte sin position som ekspert til at gøre den superviserede opmærksom på det, som han/hun definere som hensigtsmæssig adfærd hos den superviserede. Det kan ske gennem anvisninger, forslag eller lineære spørgsmål, som skal føre den superviserede frem til de nødvendige forståelser af, hvad der er defineret som hensigtsmæssigt af supervisor.

Det kræver selvsagt at eksperten(supervisor) er meget afklaret og sikker på, hvad der kan defineres som hensigtsmæssigt i situationen både ift. egen viden som ekspert og i forhold til den kontekst situationen udspiller sig i. Det forudsætter også en høj grad af tillid fra den superviserede side på denne måde at tage imod gode råd og anvisninger.

Når supervisor optræder som vejsøgende sammen med den superviserede, vil han/hun stille spørgsmål, som forsøger at udfolde situationens kompleksitet og forsøge at inddrage den superviseredes erfaringer, perspektiver og holdninger i refleksionerne over eller håndteringen situationerne. Denne form for supervision kan ofte rumme en højere grad af kompleksitet og adressere spørgsmål, som ikke nødvendigvis kan tolkes entydigt. Samtidigt kan denne form for supervision også opleves frustrerende for den superviserede, hvis han/hun gerne vil have håndgribelige og umiddelbart virksomme løsningsforslag.

Supervision i NySeSam

I Sesam-LS defineres du som 'koshi' (supervisor), som havende omfattende viden på det faglige og fagdidaktiske område som undervisningen drejer sig om. Supervisor antages at have et godt overblik over litteratur med relation til undervisning og læring af stoffet, samt have overværet undervisning i mange forskellige klasserum. Supervisor bruger overvejende sin viden til at være vejsøgende sammen med supervisanten.

Samtidigt er supervisor eksponent for nogle grundtanker om hensigtsmæssig levering af undervisningen, som tidligere nævnt under afsnittet om den implicitte normative dimmension. Det kan blandt andet få betydning for forventningerne fra de superviserede, og det betyder også, at der ikke kan superviseres på hvad som helst, uanset om det kunne være pædagogisk eller didaktisk relevant. (fx vil supervisor i NySeSam sjældent pege på gruppesammensætninger som årsag/forklaring eller ændringsmulighed. (Det antages at her er supervisanten suveræn ekspert))

Supervisor i NySeSam er overvejende:

- Nysgerrigt medundersøgende
- Konkluderende ud fra konkrete observationer
- Vejsøgende

Rollefordeling

Husk at forholdet er et asymmetrisk forhold, men asymmetrien søges udtonet. Din rolle er at være den superviseredes medundersøger. At hjælpe med at se det generelle i den specifikke konkrete undervisning. At bidrage med et bredere perspektiv på observeret elevlæring, og bidrage med forslag til hvordan læringen kan søges optimeret, og hjælpe med hvilke konklusioner om elevlæring (og undersøgelsesformålet) der kan drages på baggrund af den sete undervisning.

Det er også væsentligt at være opmærksom på, at den superviserede sætter sig selv på spil og blotter sig selv i sin praksis og ved at invitere et fremmed blik ind i denne praksis med henblik på at udvikle denne praksis,

Tavs og eksplicit viden

En væsentlig udfordring i forbindelse med supervisionen i dette projekt bliver at få etableret et meningsfuldt refleksionsrum for de superviserede, hvor NySeSam som ramme møder de superviseredes personlige erfaringer og lokalt situerede teorier om, hvad der er kvalitet i undervisningen.

Det er også en væsentlig pointe, at en del af den superviseredes viden og færdigheder i forhold til at undervise og agere i den didaktiske praksis fremstår som tavs viden og ikke nødvendigvis umiddelbart kan italesættes og ekspliciteres. Eksempelvis kan læreren velmenende have en tendens til at give lange forklaringer for eleverne, ud fra den ubevidste overbevisning, at det kan få eleverne til at opnå indsigt. Det kan også ses at undervisningen er særdeles velfungerende, dvs. eleverne lærer tydeligvis det intenderede, men læreren er ude af stand til at pege på de virksomme delelementer i undervisningen.

Mødet mellem supervisor og den superviserede inkluderer altså et blik for, at den superviserede i forhold til nogle dele af egen praksis, kan beskrive og sætte begreber på praksis, mens andre dele er svære at sprogliggøre og italesætte, men at det ikke betyder, at han/hun ikke har kompetencer på området.

Det betyder også, at der må afsættes tid til, og at det kræver tålmodighed, at gøre tavs viden eksplicit.

Inddragelse af kontekst i øvrigt

Det kan være svært at afgøre i hvilket omfang, at det er aktuelt og hensigtsmæssigt at inddrage forhold, som ikke fremgår af videoobservationen. Det kunne f.eks. være:

- elevers forudsætninger og baggrund
- hændelser før og efter
- sammenhæng med andre undervisningsgange

I SeSam-LS vil det første og sidste punkt være indtænkt i undervisningsplanen, hvorimod tilfældige hændelser ofte ikke vil være relevant at inddrage.

Det er et problem hvis konteksten først og fremmest bliver brugt til at retfærdiggøre og fastholde eksisterende forestillinger om elevers viden frem for at åbne for tilegnelsen af nye indsigter. Det kan være relevant at inddrage konteksten, men det er vigtigt at fastholde, at det I har fælles foran jer er videooptagelsen, således at analysen af lektionen ikke bliver spekulativ eller båret af bekvemmelighed.

Hvad er det vigtigt at italesætte over for den superviserede inden supervisionen?

Her er en række anbefalinger til, hvad der bør italesættes op til supervisionen:

- Lektionsstudier som fremstillet her i NySeSam fanger ikke alt, som har betydning pædagogisk og didaktisk, og der kan være langt fra studielektionen til den superviseredes undervisning i almindelighed. Men de kan hjælpe til at få øje på interessante forhold i undervisningen og fastholde opmærksomheder, som det er interessant at arbejde videre med.
- Man kan være uenig i det fokus NySeSam anlægger, men i denne proces er det vigtigt, at det er det, man tager det konceptuelle setup for gode varer og bruger det som afsæt for samtalerne om undervisningen og læringen.
- Det interessante i selve undervisningsplanen er ikke tidsangivelser. Det interessante er overvejelserne over hvad eleverne forventes at kunne og gøre. Udfærdigelse af planen skulle gerne hjælpe til at træffe valg ift. undersøgelsesformålet og hvad man vælger at arbejde med at blive bedre til, eller klogere på i undervisningen
- Det centrale er altså den fælles interesse i at blive klogere på forbindelsen mellem intenderet undervisning og elevernes læring.

Fra observation til supervision - hvad gør supervisanten og supervisor

Supervisanten

1. Supervisanten afprøver udstyret uden for undervisningen, hvor der ikke er forhold, som kan forstyrre, stresse eller distrahere. Her optages en kort sekvens (indholdet er fuldstændigt irrelevant), som uploades og deles med supervisor. Alt sammen for at lære udstyret at kende og sikre, at der er hul igennem.
2. Supervisanten udtænker en kommende lektion med et undersøgelsesformål og nedskriver planen i det korte to-siders format. Og sender denne pr. mail til supervisor senest to dage før aftalt samtale (45 minutter)

3. Supervisanten afholder 30 minutters samtale med supervisor om lektionsplanen (planlægningen). Og bruger efterfølgende, forud for afholdelsen af lektionen 15 minutter på at revidere planen.
4. Supervisanten optager en lektionen på 45 minutter
5. Efterfølgende sendes (uploades) det optagede til supervisor sammen med den reviderede og faktuelt anvendte plan.
6. Supervisanten ser/genser video forud for 30 minutters samtale med supervisor

Supervisor

1. Supervisor modtager lektionsplanen hhv. videoen. Supervisor bekræfter modtagelsen og den aftalte tid for samtalerne.
2. Supervisor gennemlæser plan og forbereder sig på at formulere relevant feedback på planen
    a. Der vejledes på elevperspektivet: er det tydeligt hvordan lektionen kommer til at forløbe set fra et elevsynspunkt?
    b. Der vejledes med eksisterende viden om undersøgelsesformålet
    c. Der udvises interesse for det valgte undersøgelsesformål, og der vejledes i at tydeliggøre og bruge fagdidaktiske virkemidler.
3. Supervisor gennemser video og forbereder sig på at give feedback ud fra konkrete observationer i videoen. Supervisor overvejer tentative konklusioner i forhold til, hvad lektionen viste, primært i forhold til undersøgelsesformålet, sekundært andre forhold som uventet viste sig at have indvirken på lektionen, eller særligt tydeligt viste elevlæring.

Supervisor og supervisant

1. Supervisor og supervisant taler sammen via telefon. De har begge lektionsplanen hhv. videoobservationen åben på en computer med mulighed for at spole til relevante episoder, som kan man kan reflektere over.
2. Supervisor lægger ud ved at rammesætte hver samtale. Herefter kan der være plads til umiddelbare reaktioner eller afklarende spørgsmål fra supervisanten. Supervisor styrer tiden, og sørger for at samtalen holdes på sporet og fx ikke digresserer ud i samtale om specifikke pædagogiske forhold for enkeltelever.
3. Samtale om planen: Efter rammesætning forsøger supervisor kort at fortælle hvorledes han/hun forstår undersøgelsesformålet og gangen i lektionen og hvordan de enkelte aktiviteter skal føre frem til den intenderede læring hos eleverne. Derefter glider samtalen over i dialog hvor supervisor og supervisant i fællesskab søger at gøre lektionen skarpere, mere velfungerende/sammenhængende og mere egnet til at belyse undersøgelsesformålet. Samtalen afrundes med enighed om ting der skal revideres forud for lektionens afholdelse.
4. Samtale om video: Efter rammesætning lægger supervisanten ud med sin egen korte redegørelse for lektionens forløb i forhold til det planlagte og i forhold til undersøgelsesformålet. (5min) Derefter gør supervisor på baggrund af sine konkrete observationer rede for hvad lektionen viste (5min). Disse to første indlæg gennemføreres uden indblanding fra den anden. Derefter overgår samtalen til en dialog.

Til sidst bruges der tid på at supervisor opsummerer de væsentligste fælles konklusioner fra drøftelsen/dialogen.

Eksempler på typer af feedback på observationerne

Det ofte være en fordel, at der både indgår vejvisende og vejsøgende elementer i supervisionen. De vejvisende elementer kan tydeliggøre de forventninger fra dig som supervisor og de forventninger, som ligger implicit i NySeSam. De vejsøgende elementer kan understøtte supervisantens oplevelse af sig selv, som kompetent og myndig ift. undersøgelsen af undervisningen, elevernes læring og styrke oplevelsen af ejerskab.

Jf. tidligere afsnit er Vejvisende kendetegnet ved at være lineær og med en tydelig umiddelbar kobling mellem årsag og virkning. Vejsøgende er kompleks og med en mere diskuterende og inddragende tilgang til at reflektere over sammenhænge og relationer med superviserede.

I NySeSam vil supervisonen være forskudt mod det vejsøgende.

Feedback er herunder inddelt i eksempler på vejvisende og vejsøgende spørgsmål og tilbagemeldinger i forhold til hhv. lektionsplanen og video-observationen.

Supervision på lektionsplanen:

Vejvisende:

”Planen ser fin ud, men jeg foreslår at klassen skal diskutere et fælles eksempel mere inden grupperne går i gang. Til gengæld kan vi droppe introduktionen af CL-strukturen og lade dem arbejde som de selv vælger.”

”Det er dejligt at se der er lagt op til en god afveksling mellem der hvor læreren siger noget og hvor eleverne siger noget. Jeg tænker dog ikke vi skal starte med den gennemgang, da det vil fjerne begrundelsen af elevernes egen undersøgelse. Der skal være noget de reelt skal finde ud af. Noget de ikke for fortalt på forhånd.”

”Vi skal prøve at forestille os hvad eleverne vil svare i konkret faglig henseende. Ikke blot at nogen siger ’jeg kan ikke finde ud af det, osv.’ Jeg forstiller mig nogle vil lægge tilbudspriserne sammen inden de beregner momsen.

Vejsøgende:

”Jeg synes det ser spændende ud at bruge GeoGebra til denne del aktiviteten, men jeg er ikke helt med på hvad eleverne egentlig skal lave. Kan du fortælle lidt mere om det?”

”Når undersøgelsesformålet er at blive klogere på elevernes ræsonnementer om chance er jeg usikker på om det er en god ide at lade dem analysere et spil de kender. Hvad ville der ske ved at de så på et ukendt, fx på Efrons terninger?”

”Det er dejligt at du har tænkt på så mange forskellige løsningsmuligheder. Men hvilke er egentlig de simpleste? Hvad for nogle tror du eleverne først finder frem til?”

”Dine forslag til lærerstøtte er meget opfindsomme. Lad os snakke lidt mere om hvilken retning vi tror eleven (tankemæssigt) vil gå i når de får de forskellige hints.”

Supervision på video-observation:

Vejvisende:

”Jeg lagde mærke til at elevere ved de fire af bordene alle kom med argumenter der havde med omvendt proportionalitet at gøre, fx sagde elev x at … og elev y sagde … og det var jo også dem som havde været i værksted A. Derfor synes jeg det er tydeligt at den aktivitet er bedre end den i værksted B”

”Det var sjovt at se hvordan de foretog opmålingen af figuren. Der var faktisk flest der gjorde det på x måden end y måden. Det er godt at vide til en anden gang, men der var også flere andre måder vi ikke havde forestillet os. Fx har vi set at det også kan gøres som Sofie gjorde. Hvad var der godt ved hendes måde?”

”Jeg er menig med dig, når du siger at overgangen fra den første aktivitet til den anden aktivitet ikke ser ud til at fungere. Og du kommer til at skulle hjælpe parrene alt for meget. Jeg tror det er fordi du/vi får sagt ’beregne’ og det gør, at meget få prøver at danne sig et overblik med en tegning eller lignende inden de går i gang.”

Vejsøgende:

”Der kom nogle meget alternative forslag til arealformler fra gruppe 2, (vi kan se det på videoen ved tid 15:24) Kan du fortælle lidt mere om, hvad det er de siger til dig da du kommer derned?”

”Det er tydeligt at eleverne ved stationen hvor de selv skulle finde frem til rumfanget af den underlige flaske arbejdede meget mere koncentreret end der hvor de skulle afprøve de givne formel-forslag. Hvad mener du årsagen kan være, og hvad lagde du særlig mærke til i snakken ved de enkelte stationer?” … ”Jeg tror dette viser os noget vigtigt om elevernes læring i forbindelse med modellering jf. vores undersøgelsesformål”

”Klassesamtalen efter eleverne selv havde arbejdet gik bare fantastisk, hva’! Hvis vi ser i tidsrummet fra 30:20 til 42:50 så er det opløftende hvordan eleverne bygger oven på, og bruger, hinandens forklaringer. Hvad mener du er årsagen til at det gik så godt? Og hvorfor lod du eleverne komme til orde i netop dén rækkefølge?”

”Du siger i din indledning at du selv er utilfreds med evalueringen til sidst i timen. Hvis vi nu ser på situationen der begynder ca. 41.30, hvad mener du om de spørgsmål du får stillet til eleverne? … jeg kan se på videoen at du er overrasket over elevens svar… Hvordan kan vi stille spørgsmål som faktisk får eleverne til at give os noget brugbar information om deres læring?”

Afrunding

At ændre fagdidaktisk praksis er ikke ukompliceret. Heller ikke bare fordi man har et koncept at læne sig op ad. I sidste ende er man som supervisor afhængig af den superviseredes lyst og evne til at udfordre sin egen praksis. Grundlæggende kan man sammen med den superviserede prøve at få øje på forhold, som er interessante og invitere til enten en dybere forståelse af, hvad det er, som sker, og/eller som inviterer til at man forsøger at ændre i praksis og prøve nye ting af. Der findes sjældent garantier i fagdidaktisk praksis, men hvis man med nysgerrighed og opfindsomhed er villig til at afprøve nye tiltag, øger man chancen for at blive klogere og mere kompetent. SeSam-LS er ikke en facitliste, men kan hjælpe med en systematisk tilgang til at udvikle egen undervisning og viden om elevernes læring.

9.6 Appendix F: Information about video recording and data collection

Institut for Læreruddannelse

KØBENHAVNS PROFESSIONSHØJSKOLE

NySeSam i Matematik – NYsgerrigt Ser vi undervisning SAMmen

Information om videooptagelse mv. til forskningsformål - elever/forældre

August 2024

Københavns Professionshøjskole

Læreruddannelsen
Institut for Læreruddannelse
Det Lærerfaglige Fakultet
Humletorvet 3
1799 København V
Tlf. nr. 70 89 09 90
info@kp.dk
www.kp.dk
Sagsnr.

Videooptagelsens formål
I SeSam optages video af undervisning som bruges af klassens lærer og en "SeSam"-partner fra Københavns Professionshøjskole til at samtale om udvikling af læring og undervisning. Formålet er at fastholde undervisningssituationer, så der kan drages fælles konklusioner om sammenhængen mellem undervisning og læring.

Projektansvarlige
Camilla Hellsten Østergaard, lektor, ph.d., Københavns Professionshøjskole

Fortrolighed
Vi behandler og opbevarer naturligvis videooptagelserne fortroligt. Se evt. mere på https://www.kp.dk/om+os/databeskyttelsespolitik

Hjemmelsgrundlag
Vi indsamler og behandler børnenes personoplysninger til videnskabeligt og/eller statistisk formål. Som uddannelses- og forskningsinstitution har vi en forpligtelse og generel hjemmel til at behandle personoplysninger til det formål.
Vi indsamler og behandler personoplysninger med hjemmel i Databeskyttelsesforordningens artikel 6 stk. 1 litra e.
Videooptagelserne opbevares af Københavns Professionshøjskole 2 år efter projektets afslutning, 31.12.2026, hvorefter videoerne vil blive slettet.

Frivillighed
Som elev eller forælder/værge man kan til enhver tid ønske ikke at indgå i videomateriale eller testindhentning. Det gøres ved at sende en mail til choe@kp.dk

Klageadgang
Hvis du ønsker at klage over vores behandling af dine personoplysninger, kan du rette henvendelse til:
DPO: Karsten Holt, DPO-Team, telefon: 71 74 90 20,
e-mail: kp@dpoteam.dk
eller
Datatilsynet, Carl Jacobsens Vej 35, 2500 Valby,
Telefon: 33 19 32 00, e-mail: dt@datatilsynet.dk

Side 1 af 2

Institut for Læreruddannelse, Københavns Professionshøjskole

Med venlig hilsen, på vegne af alle i projektgruppen,

Camilla Hellsten Østergaard
Lektor og ph.d.
Københavns Professionshøjskole
choe@kp.dk
Tlf. 26 256 257

9.7 Appendix G: Information letter for potential participants

KP | Københavns Professionshøjskole

SeSam – at Se Sammen på undervisningen og læringen – Pilot 1.- 6. klassetrin, Matematik

Kære potentielle deltager, lærer - skoleleder - skolechef

Der er mange undersøgelser som tyder på, at når man drøfter sin undervisning med andre, så udvikler man sig selv, udvikler undervisningen, får en større tilfredsstillelse ved egen lærergerning og eleverne lærer mere. Vi er interesseret i at finde ud af, hvordan vi bedst gør dette - sammen.

Hvad får du og din skole ud af at være med i SeSam? Som lærer får du et forløb til professionel udvikling, der har fokus på netop din undervisning og din omgang med dine elever og deres læring. Det er en hjælp til at se på din undervisning, interaktionen med eleverne og deres læring med et mere udefra kommende blik, som er svært at have når man står midt i undervisningens hektiske handletvang. Sparring og samtale om de udfordringer du selv møder i dit undervisningsarbejde.

SeSam fungerer i korte træk på den måde, at man som lærer mødes med en særligt uddannet "SeSam-kollega" fra Københavns Professionshøjskole til to introducerende kursusgange. Imellem og i forlængelse af disse, får du et specielt kamera til at videooptage noget af din undervisning, som du efterfølgende ser på sammen med din SeSam-kollega. Videoerne udveksles over nettet, ligesom samtalen omkring dem. I samarbejder om at blive klogere på elevernes læring og interaktionen i klasserummet på en systematisk måde.

SeSam drejer sig om undervisning og i sidste ende om at gøre læringen bedre hos eleverne. Derfor vil eleverne i løbet af projektet skulle tage tre tests af en times varighed.

Hvis du/I vil være med, vil der naturligvis skulle bruges lidt tid på det. Som lærer kan man forvente at bruge ca. 27 timer på deltagelsen hen over et skoleår. 2 gange 3 timer til de introducerende kursusgange (Hertil kan der komme tid til transport), 18 timer til selve det videostøttede samtaleforløb og endelig 3 timer til interviews, spørgeskemabesvarelse mv. Skolen kompenseres økonomisk for lærerens deltagelse. 27 timer á 283,50 kr.

Projektet stiller gratis kursusgangene og SeSam-kollegaen til rådighed. Desuden må videokameraet efterfølgende beholdes af skolen/deltageren.

Du kan tilmelde dig ved at skrive en mail til Camilla H. Østergaard choe@kp.dk

Bedste hilsner

Julie Vangsøe og Camilla H. Østergaard

Københavns Professionshøjskole

9.8 Appendix H: Manual for administering test materials at schools

Vejledning Matematikprofilen 1.-6. klasse

I NySeSam er testmaterialet baseret på Matematikprofilen. Nærmere beskrivelser af materialet og tankerne bag kan findes på https://matematikprofilen.gyldendal.dk/.

Oprindeligt består Matematikprofilen af tre dele. Opgaverne i del A er multiple choice og tager udgangspunkt i stofområder. Opgaverne i del B tager udgangspunkt i stofområder og kompetencer og opgaverne i del C tager udgangspunkt i kompetencerne. I B- og C-opgaverne skal eleverne tegne og/eller skrive deres svar i de enkelte rammer.

I dette testmateriale er A-, B- og C-opgaver fra Matematikprofilen blevet opdelt i tre tests, således at alle tre typer af opgaver er at finde i alle tre deltests. De tre deltests er navngivet XA, XB og XC [X for klassetrin] (A, B og C ej at forveksle med de tre forskellige opgavetyper fra det oprindelige Matematikprofilen-materiale). XA gennemføres først, efterfulgt af tre måneders undervisning, som den normalt finder sted på skolen. Efter ca. 3 måneder gennemføres XB, hvilket efterfølges af tre måneder, hvor læreren får sparring fra KP på sin undervisning. Efter disse tre måneder gennemføres deltest XC.

Hver test forventes at kunne gennemføres på en time. Nogle elever vil muligvis være færdige efter 20 min., mens andre har brug for ekstra tid. Det er vigtigt, at eleven forsøger at besvare alle spørgsmål. I 1.-3. klasse har eleverne muligvis brug for en pause efter 15-20 min. arbejde. På alle klassetrin er det op til læreren at bedømme, om der er brug for pause til enten alle eller udvalgte elever. Der er også plads på nogle af siderne til, at eleverne kan tegne, hvis de har brug for et afbræk. De elever, der er hurtigt færdige kan læse, tegne el.lign., mens resten af klassen bliver færdige. Det kan være en god ide, at de har bøger/papir/penalhus klar ved bordet, så de ikke kommer til at forstyrre de andre elever.

For alle klassetrin gælder det, at eleverne må få hjælp til at læse og forstå opgaverne, så det ikke er eventuelle læseproblemer, der gør, at de ikke kan løse opgaverne. De må dog ikke få hjælp til at løse dem. Herunder må de heller ikke få hjælp til at vælge en strategi eller løsningsmetode. I opgaver, som evaluerer om eleverne forstår et bestemt fagord, må læreren ikke forklare fagordets betydning for eleverne. Eventuelle opklarende spørgsmål fra eleverne vedrørende opgaveinstruktionen eller de tilknyttede illustrationer, som ikke direkte omhandler selve besvarelsen af opgaven, må gerne besvares. *Du kan vælge, at klassen arbejder samtidigt med opgaverne, og at du løbende læser en opgave højt og giver eleverne et passende tidsrum til at besvare hver opgave. En anden mulighed er, at du går rundt blandt eleverne og læser højt for de elever, der har brug for det.*

Eleverne må gerne bruge de hjælpemidler, som de er vant til fra den daglige undervisning som fx taltavle og forskellige tællematerialer, men ikke lommeregner, med mindre det står nævnt direkte. Eleverne må også bruge papir til noter og udregninger. Disse noter må ikke viskes ud, og skal afleveres sammen med testen, da de tæller med i vurderingen. *Nogle af opgaverne i del B og del C* *kan eleverne ikke løse uden bestemte materialer.* Hvilke materialer det drejer sig om, står anført under beskrivelsen til de enkelte klassetrin.

Det er hensigten, at eleverne skriver evt. egne noter eller laver tegninger, der kan støtte deres tænkning, direkte på opgavearket. Disse noter eller tegninger indgår i vurderingen af elevernes arbejde sammen med elevernes løsningsforslag i flere af opgaverne og i den efterfølgende formative evaluering. Det er derfor vigtigt, at eleverne ikke visker deres noter og tegninger ud.

Når hele klassen er færdig med testen, må læreren meget gerne samle dem ind på en måde, så tests gennemført af elever, der har siddet ved siden af hinanden, ligger efter hinanden, da det gør det lettere for dem, der retter testen at se, hvis nogle elever har arbejdet sammen.

1. klasse

Testens forside udfyldes på forhånd af klassens matematiklærer eller skolens matematikvejleder.

Det er en god ide at have to lærere i klassen, dels da der er brug for hjælp til oplæsning og dels da der er opgaver, hvor besvarelsen er en mundtlig besvarelse (gælder kun 1C).

Da de fleste elever vil have brug for hjælp til at læse opgaverne, kan læreren overveje at læse opgaverne op for hele klassen på en gang. Sæt god tid mellem oplæsning, så alle elever kan nå at besvare opgaven. Opgaver fra del A, som er multiple choice, kræver ikke så lang tid, til gengæld kan det være svært for eleverne at overskue de forskellige muligheder ved fælles oplæsning. Alternativt kan klassen besvare testen i mindre grupper, så der sidder en lærer sammen med dem, der kan hjælpe til med at læse og forstå opgaverne.

- 1A: Alle elever skal have centicubes (fire til hvert barn, opgave 12 side 5) og lineal (opgave 13 side 5)
- 1B: Alle børn skal have centicubes (25 til hvert barn, eller til deling ved bordet, opgave 9 side 5)
- 1C: De to opgaver på side 7 og 8 skal besvares mundtligt. Det vil derfor være nødvendigt at én lære kan tage dialogen med et barn ad gangen fra lektionens start. Læreren skal skrive noter i hver enkelt test til denne opgave, og skal så vidt muligt ikke tolke på elevernes svar, men blot gengive elevels besvarelse så præcist som muligt.

Eleverne skal bruge lommeregner i opgave 4, men må ellers ikke bruge den

Information til eleverne:

Opgavernes numre kommer i en underlig rækkefølge – I skal ikke tage jer af, hvilke numre, der står i øverste i toppen af hver opgave. Nummeret skal kun bruges af den person, der retter opgaverne bagefter.

Nogle af opgaverne vil være svære, men I skal bare svare så godt i kan. Det er bedre, at I skriver/tegner, hvordan I tror den skal løses, end at I slet ikke svarer.

Ved alle opgaver er der plads til, at I kan skrive mellemregninger eller tegne illustrationer. Hvis der er opgaver, I ikke forstår, så ræk hånden op, så kan I få hjælp til at læse eller forstå dem, men I kan ikke få hjælp til at løse opgaverne.

Når I er færdige med opgaverne, skal I lade dem ligge på bordet, så samler jeg dem ind senere.

2. klasse

Testens forside udfyldes på forhånd af klassens matematiklærer eller skolens matematikvejleder.

Det er en god ide at have to lærere i klassen, dels da der er brug for hjælp til oplæsning og dels da der er opgaver, hvor besvarelsen er en mundtlig besvarelse (gælder kun 2B).

Da de fleste elever vil have brug for hjælp til at læse opgaverne, kan læreren overveje at læse opgaverne op for hele klassen på en gang. Sæt god tid mellem oplæsning, så alle elever kan nå at besvare opgaven. Opgaver fra del A, som er multiple choice, kræver ikke så lang tid, til gengæld kan det være svært for eleverne at overskue de forskellige muligheder ved fælles oplæsning. Alternativt kan klassen besvare testen i mindre grupper, så der sidder en lærer sammen med dem, der kan hjælpe til med at læse og forstå opgaverne.

Eleverne må ikke bruge lommeregner. De må bruge de hjælpemidler, de er vant til fra undervisningen. Det kan fx være lineal, legepenge, centicubes eller andre tællematerialer.

- 2A
    - Eleverne skal bruge en lineal (opgave 12 side 4). Hver elev skal bruge 5 centicubes og evt. et spejl (opgave 5 side 7).
- 2B
    - Talkort 1-5 til alle børn (opgave 1 side 6).
    - I opgave 6 (side 8) skal eleven være i dialog med læreren. Det vil derfor være nødvendigt, at én lærer kan tage dialogen med et barn ad gangen fra lektionens start. Læreren skal skrive noter i hver enkelt test til denne opgave, og skal så vidt muligt ikke tolke på elevernes svar, men blot gengive elevels besvarelse så præcist som muligt.
- 2C
    - Centicubes 12 (eller 24) til hvert barn (opgave 10 side 5).

Information til eleverne:

Opgavernes numre kommer i en underlig rækkefølge – I skal ikke tage jer af, hvilke numre, der står i øverste i toppen af hver opgave. Nummeret skal kun bruges af den person, der retter opgaverne bagefter.

Nogle af opgaverne vil være svære, men I skal bare svare så godt i kan. Det er bedre, at I skriver/tegner, hvordan I tror den skal løses, end at I slet ikke svarer.

Ved alle opgaver er der plads til, at I kan skrive mellemregninger eller tegne illustrationer. Hvis der er opgaver, I ikke forstår, så ræk hånden op, så kan I få hjælp til at læse eller forstå dem, men I kan ikke få hjælp til at løse opgaverne.

Når I er færdige med opgaverne, skal I lade dem ligge på bordet, så samler jeg dem ind senere.

3. klasse

Testens forside udfyldes på forhånd af klassens matematiklærer eller skolens matematikvejleder.

Det er en god ide at have to lærere i klassen, dels da der er brug for hjælp til oplæsning og dels da der er opgaver, hvor besvarelsen er en mundtlig besvarelse (gælder kun 3A).

Da de fleste elever vil have brug for hjælp til at læse opgaverne, kan læreren overveje at læse opgaverne op for hele klassen på en gang. Sæt god tid mellem oplæsning, så alle elever kan nå at besvare opgaven. Opgaver fra del A, som er multiple-choice, kræver ikke så lang tid, til gengæld kan det være svært for eleverne at overskue de forskellige muligheder ved fælles oplæsning. Alternativt kan klassen besvare testen i mindre grupper, så der sidder en lærer sammen med dem, der kan hjælpe til med at læse og forstå opgaverne.

Eleverne må ikke bruge lommeregner. De må bruge de hjælpemidler, de er vant til fra undervisningen. Det kan fx være lineal, legepenge, centicubes eller andre tællematerialer.

- 3A
    - Hver elev skal have 3 centicubes (opgave 9 side 5).
    - Opgave 5 og 6 på side 7 og 8 skal besvares mundtligt. Det vil derfor være nødvendigt at én lære kan tage dialogen med et barn ad gangen fra lektionens start. Læreren skal skrive noter i hver enkelt test til denne opgave, og skal så vidt muligt ikke tolke på elevernes svar, men blot gengive elevels besvarelse så præcist som muligt.
    - [I kan evt. benytte 3B som den første test, så elever og lærere er bekendte med materialet inden disse opgaver kommer i spil.]
- 3C
    - Eleverne skal benytte et geometriprogram, fx GeoGebra, og til sidst printe opgaverne ud og vedlægge dem (opg. 4 side 8).

Information til eleverne:

Opgavernes numre kommer i en underlig rækkefølge – I skal ikke tage jer af, hvilke numre, der står i øverste i toppen af hver opgave. Nummeret skal kun bruges af den person, der retter opgaverne bagefter.

Nogle af opgaverne vil være svære, men I skal bare svare så godt i kan. Det er bedre, at I skriver/tegner, hvordan I tror den skal løses, end at I slet ikke svarer.

Ved alle opgaver er der plads til, at I kan skrive mellemregninger eller tegne illustrationer. Hvis der er opgaver, I ikke forstår, så ræk hånden op, så kan I få hjælp til at læse eller forstå dem, men I kan ikke få hjælp til at løse opgaverne.

Når I er færdige med opgaverne, skal I lade dem ligge på bordet, så samler jeg dem ind senere.

4. klasse

Testens forside kan udfyldes på forhånd af klassens matematiklærer eller skolens matematikvejleder.

Det er en god ide at have to lærere i klassen, da der er brug for hjælp til oplæsning.

Eleverne må bruge de hjælpemidler, de er vant til fra undervisningen. Det kan fx være lineal, centicubes eller andre tællematerialer og lommeregner. Dog ikke lommeregner i multiple-choice-opgaverne.

- 4A: Eleverne skal benytte et geometriprogram, fx GeoGebra, og til sidst tegne og forklare deres løsning (opgave 6 side 8). Det kan være en fordel at lade en lille gruppe elever ad gangen arbejde med opgave 6, så du undervejs kan skrive noter om, og hvordan de bruger geometriprogrammet.

Information til eleverne:

Opgavernes numre kommer i en underlig rækkefølge – I skal ikke tage jer af, hvilke numre, der står i øverste i toppen af hver opgave. Nummeret skal kun bruges af den person, der retter opgaverne bagefter.

Nogle af opgaverne vil være svære, men I skal bare svare så godt i kan. Det er bedre, at I skriver/tegner, hvordan I tror den skal løses, end at I slet ikke svarer.

Ved alle opgaver er der plads til, at I kan skrive mellemregninger eller tegne illustrationer. Hvis der er opgaver, I ikke forstår, så ræk hånden op, så kan I få hjælp til at læse eller forstå dem, men I kan ikke få hjælp til at løse opgaverne.

Når I er færdige med opgaverne, skal I lade dem ligge på bordet, så samler jeg dem ind senere.

5. klasse

Testens forside kan udfyldes på forhånd af klassens matematiklærer eller skolens matematikvejleder.

Det er en god ide at have to lærere i klassen, da der er brug for hjælp til oplæsning.

Eleverne må bruge de hjælpemidler, de er vant til fra undervisningen. Det kan fx være lineal, centicubes eller andre tællematerialer og lommeregner. Dog ikke lommeregner i multiple-choice-opgaverne.

- 5A: I opgave 13 (side 5) skal eleverne bruge en vinkelmåler.
- 5B: I opgave 6 (side 8) skal eleverne arbejde i et regneark, og til sidst skrive deres løsning i tabellen i hæftet. Det kan være en fordel at lade en lille gruppe elever ad gangen arbejde med opgave 6, så du undervejs kan skrive noter om, hvordan de bruger regnearket.
- 5C

Information til eleverne:

Opgavernes numre kommer i en underlig rækkefølge – I skal ikke tage jer af, hvilke numre, der står i øverste i toppen af hver opgave. Nummeret skal kun bruges af den person, der retter opgaverne bagefter.

Nogle af opgaverne vil være svære, men I skal bare svare så godt i kan. Det er bedre, at I skriver/tegner, hvordan I tror den skal løses, end at I slet ikke svarer.

Ved alle opgaver er der plads til, at I kan skrive mellemregninger eller tegne illustrationer. Hvis der er opgaver, I ikke forstår, så ræk hånden op, så kan I få hjælp til at læse eller forstå dem, men I kan ikke få hjælp til at løse opgaverne.

Når I er færdige med opgaverne, skal I lade dem ligge på bordet, så samler jeg dem ind senere.

6. klasse

Testens forside kan udfyldes på forhånd af klassens matematiklærer eller skolens matematikvejleder.

Det er en god ide at have to lærere i klassen, da der er brug for hjælp til oplæsning.

Eleverne må bruge de hjælpemidler, de er vant til fra undervisningen. Det kan fx være lineal, centicubes eller andre tællematerialer og lommeregner. Dog ikke lommeregner i multiple-choice-opgaverne.

- 6A: I opgave 6 (side 8) skal eleverne arbejde med et dynamisk geometriprogram, fx GeoGebra, og til sidst tegne og forklare deres løsning i hæftet. Det kan være en fordel at lade en lille gruppe elever ad gangen arbejde med opgave 6, så du undervejs kan skrive noter om, hvordan de bruger geometriprogrammet.
- 6C: Alle eleverne skal bruge sømpapir (kopiark) til opgave 3 (side 7).

Information til eleverne:

Opgavernes numre kommer i en underlig rækkefølge – I skal ikke tage jer af, hvilke numre, der står i øverste i toppen af hver opgave. Nummeret skal kun bruges af den person, der retter opgaverne bagefter.

Nogle af opgaverne vil være svære, men I skal bare svare så godt i kan. Det er bedre, at I skriver/tegner, hvordan I tror den skal løses, end at I slet ikke svarer.

Ved alle opgaver er der plads til, at I kan skrive mellemregninger eller tegne illustrationer. Hvis der er opgaver, I ikke forstår, så ræk hånden op, så kan I få hjælp til at læse eller forstå dem, men I kan ikke få hjælp til at løse opgaverne.

Når I er færdige med opgaverne, skal I lade dem ligge på bordet, så samler jeg dem ind senere.

9.9 Appendix I: Statistical analysis of test data

Om anvendelse af MatematikProfilen til gentagne målinger af færdigheder i matematik

1. Introduktion

Der er flere forskellige grunde til at bruge gentagne målinger ved hjælp af pædagogiske test til at vurdere progressionen af elevers færdigheder på forskellige tidspunkter af et skoleår. Den første man formodentlig tænker på er forbindelse med evaluering af pædagogiske forsøgsarbejder. En anden grund til at anvende gentagne test er at evaluere effekten af intensive kurser for en større eller mindre gruppe af elever med faglige problemer. I begge disse tilfælde er interessen ikke kun er at afprøve om der er en statistisk signifikant effekt af forsøgsarbejdet eller kurset, men også at vurdere om effekten modsvarer det arbejde og de ressourcer, der er lagt i arbejdet. Den tredje og lige så vigtige årsag er at lærere kan have behov og interesse i at følge op på enkelte elevers eller enkelte klassers udvikling i forbindelse med undervisningen. Et behov, der kun i ringe grad handler om statistisk signifikans, men hvor det er vigtigt at testresultaterne både fortæller læreren noget om hvad eleverne ikke kan, men også hvad de kan.

I alle disse tilfælde er det afgørende, at det er muligt at foretage gentagne målinger af elevers færdigheder på en måde, der lever op til tre krav.

- Det første er, at gentagne målinger er sammenlignelige i den forstand et testresultaterne består af værdier på en og samme færdighedsskala.
- Det andet er resultater af tidligere test ikke i sig selv har nogen effekt på de efterfølgende resultater.
- Det andet er, at der foreligger kvalitative kriteriebaserede tolkninger der forklarer hvad testresultaterne fortæller om elevers færdigheder

Det første krav synes at lægge op til at man anvender en og samme pædagogiske test hver gang elevernes færdigheder skal vurderes.

Det andet krav betyder til gengæld at konventionelle lineære tests ikke kan anvendes, fordi det er de samme opgaver eleverne bliver bedt om at bevare, hver gang der skal testes. Såkaldte adaptive tests, hvor eleverne svarer på forskellige opgaver, ville være en måde at kommer uden om det problem på, men erfaringer med bl.a. de såkaldte nationale test har vist at adaptive test ikke er praktisk anvendelige for lærere på elev eller klasseniveau.

Alternativet er derfor at udvikle flere forskellige test, hvor man i forbindelse med afprøvningen af testene sikrer sig at testresultater er kalibrerede til en og samme færdighedsskala. En opgave, der både tidsmæssigt og ressourcemæssigt kan være særdeles udfordrende. I SeSam har man forsøgt at skyde genvej ved at definere tre forskellige deltest bestående af opgaver fra Gyldendals Matematikprofil, hvor sværhedsgraderne er defineret på en sådan måde at testresultaterne kan kalibreres til de samme færdighedsskalaer, testresultaterne fra MatematikProfilens måles på, og hvor der foreligger kriteriebaserede tolkninger af testresultaterne, der gør det muligt at beskrive effekten af forsøgsarbejde og undervisning på en måde der giver mening i forhold til MatematikProfilens faglige referenceramme.

Formålet med dette bilag er at beskrive, hvorledes disse deltest er defineret og hvorledes det er muligt at måle og beskrive færdigheder målt ved hjælp af deltestene på den samme skala som MatematikProfilen anvender.

2 MatematikProfilen

Afprøvningen af MatematikProfilen accepterede tilpasning mellem svarene på opgaverne og såkaldte log-lineære Rasch modeller. I de analyser, der omtales i dette notat kontrollerer vi indledningsvis at tilpasningen til Gyldendals data accepteres og anvender Gyldendals estimater af item og personparameter som udgangspunkt i forbindelse med analyserne af data fra SeSam projektet.

2.1 Log-lineære Rasch modeller

Log-lineære Rasch modeller er generaliseringer af konventionelle Rasch modeller, som tillader lokal responsafhængighed og differentiel item funktion (DIF). Modellerne antager, at svar på opgaver afhænger af fire typer af parametre.

- Person parametre der definerer et kvantitativt mål for elevernes færdigheder
- Item parametre der definerer opgavernes sværhedsgrader
- Interaktionsparametre, der definerer styrken af interaktioner mellem items og mellem items og eksogene variable.

Log-lineære Rasch modeller antager at den samlede score er statistisk sufficient for personparameteren og at interaktionerne er uniforme i den forstand at styrken af interaktionerne er uafhængig af værdien af personparameteren.

Det vil føre for vidt at gøre mere ud af modellere her. Vi henviser til www.biostat.ku.dk/DIGRAM hvor man kan downloade DIGRAM programmet, som vi har anvendt til analyserne sammen ed brugervejledninger og artikler, vor modellerne er defineret.

2.2. Rasch modellerne færdighedsskalaer

Testresultater fra Rasch modeller og log-lineære Rasch modeller måles på tre forskellige skalaer.

- Personparameteren måles på en såkaldt logit-skala, som opfattes som et interval-skala med arbitrært nulpunkt Alle monotone funktioner af denne skala kan opfattes som måle-skalaer, men det er kun lineære transformationer, der har intervalskala-egenskaber. Der findes flere forskellige måder at estimere parameteren på. Den bedste er det såkaldt vægtede likelihood estimat, som er næsten uden systematisk bias.
- Den forventede samlede score på testen er en ikke-lineære monoton funktion af personparameteren. Den observerede score på testen er et unbiased estimat af den forventede score. Mange opfatter den som en måling på en ordinal skala. Det såkaldte maksimum likelihood estimat af personparameteren er den lig med den personparameter hvor den observerede testscore er lig med den forventede testscore.

- Kategoriseringer af den samlede score eller af estimatet af personparameteren måler færdigheden på en såkaldt ordinal kategori skala. I forbindelse med Matematikprofilen er der defineret fem kriteriebaserede ordinale kategorier, med indhold, der kan fortolkes i termer knyttet til de såkaldte Fælles mål.

Da Rasch modellen logit-skala er en abstrakt skala, der kun kan fortolkes hvis man beregner sandsynligheder for svar på de enkelte opgaver. Da disse ikke er tilgængelige, rapporteres testresultater fra Matematik profilen enten som den samlede score, som kan give mening for lærere med anvendelser af testen eller – endnu bedre i form af de kriteriebaserede kategorier. I forbindelse med SeSam projektet, vil det være disse to forer for skalaer, der anvendes.

- 2.3. Om delskalaer defineret ved log-lineære Rasch modeller

Antag, at en pædagogisk test består af tre sæt items, (A1,…,Aa,B1,…,Bb,C1,…,Cc) som tilsammen passer til en log-lineær Rasch model, hvor der ikke er lokal afhængighed mellem items fra forskellige sæt. Under denne forudsætning, vil hvert af de tre sæt passe til log-lineære Rasch modeller, men den samme person parameter og med de samme item og interaktion parametre som i den samlede skala. Summen af det tre sæts af items definerer derfor tre forskellige delskalaer der kan derfor bruges til at estimere den fælles person parameter på en skala, der er automatisk kalibreret til den skala, som den estimaterne af personparameteren fra den samlede skala måles på. Og man kan estimater den forventede samlede score på testen ud fra hver af de tre delskalaer.

Problemet er naturligvis, at usikkerheden på delskalaerne og usikkerheden på estimaterne af personparameteren og usikkerheden på den forventede samlede skala er større end usikkerheden på estimatet af personparameteren ud fra den samlede skala, men opdelingen i tre lokalt uafhængige dele har en væsentlig fordel. Man kan bruge den pædagogiske test til gentagne målinger. Man tager simpelthen testen tre gange. Første gang nøjes man med at beregne den samlede A-skala og estimerer personparameteren ud fr denne delskala. Anden og tredje gang gør man det samme med henholdsvis B- og C-skalaen, hvorefter man har tre målinger af færdigheden, der kan sammenlignes fordi de lever op til ovennævnte krav fordi de tre sæt opgaver er lokalt uafhængige. Og man kan omregne de tre estimater af personparametrene til tre uafhængige estimater af den forventede score på den samlede skala, hvis det foretrækkes at buge denne skala i stedet for Rasch modellen logit skala.

Forudsætningen for at anvender testen på denne måde er naturligvis at antallet af opgaver i de tre delskalaer er så stort at delskala målingerne ikke er alt for usikre.

Det er disse betragtninger, der førte til anvendelsen af MatematikProfilen i forbindelse med SeSam. Antallet af opgaver i MatematikProfilen er relativt stort sammenlignet med andre pædagogiske test og MatematikProfilen indeholder flere såkaldte polytome opgaver. Dette – sammen med den lokal afhængighed som også bidrager til sikre målinger – betyder at det forekom realistisk at forvente at MatematikProfilen kunne definere tre delskalaer, som var sikre nok til at SeSam kunne anvende dem.

3. SeSam projektets deltest

SeSam projektet består af seks del-projekter fra første til sjette klassetrin der anvender gentagne målinger af delskalaer med opgaver fra Gyldendals Matematikprofil for de relevante klassetrin.

Hver af Matematikprofilerne består af 32 dikotome opgave, 16 trinære og 6 polytome opgaver, hvor der scores fra 0 til 3 point. Og det var ikke vanskeligt at definere tre delskalaer ud fra hver af de seks Matematikprofiler.

SeSam projektet gik imidlertid et skridt videre for at undgå at bruge den samlede Matematikprofil hver eneste gang. I stedet for at bruge den samlede matematikprofil og beregne forskellige delskalaer, definerede SeSam i stedet tre deltest, der kun indeholdt de opgaver, som definerede delskalaerne. Forventningen var at de tre forskellige deltest, ville passe til de samme log-lineære Rasch modeller som de tre delskalaer, men at der kunne forekomme situationer hvor opgavernes sværhedsgrader i begrænset omfang kunne afvige fra opgavernes sværhedsgrad fordi konteksten og test-situationen var en anden når der kun skulle svares på opgaver fra deltesten i stedet for hele MatematikProfilen.

Disse forventninger stillede ekstra krav til analyserne. For at sikre at deltestene fungerede på samme måde som delskalaerne og for at estimere eventuelle forskelle i opgavernes sværhedsgrader, har vi derfor kontrolleret om deltestene passer til log-lineære Rasch modeller med den samme grad af lokal afhængighed mellem opgaver som i delskalaerne, men hvor der kunne være en vis grad af differentiel item funktion (DIF) i deltestene i forhold til modellerne for delskalaerne.

Modellerne blev i samtlige tilfælde accepteret. Delskala modellerne passede til SeSams deltest, men der var som forventet eksempler på DIF, som gjorde kalibreringen af deltest-resultater til MatematikProfilens færdighedsskalaer til en lidt mere kompliceret totrins procedure end hvis der ikke havde været eksempler på DIF.

Det første skridt justerer den samlede score et resultat fra en SeSam-deltest til en forventet delscore på opgaverne, hvis eleven havde besvaret hele Matematik profilen, hvorefter det andet trin konverterer dette resultat ti en samlet score på hele MatematikProfilen.

Opgaverne i de 18 deltest fremgår af efterfølgende oversigter, hvor MP referer til delscores i matematik profilen. Ud for hver opgave er der markeret en label bestående af et enkelt bogstav. Disse labels benyttes af DIGRAM programmet og vil optræde i en del af det efterfølgende output fra programmet

1.-3. KLASSE

SeSam1			SeSam2			SeSam3		
SeSam1A	SeSam1 B	SeSam1C	SeSam2A	SeSam2B	SeSam2C	SeSam3A	SeSam3B	SeSam3C

MP1A	MP1 B	MP1 C	MP2 A	MP2 B	MP2 C	MP3 A	MP3 B	MP3 C
A **A2**	D **A5**	C **A4**	A **A15**	C **A17**	E **A19**	D **A19**	C **A18**	A **A9**
B **A3**	E **A6**	H **A19**	B **A16**	D **A18**	F **A20**	E **A20**	F **A21**	B **A17**
J **A21**	F **A17**	M **A24**	G **A21**	J **A24**	H **A22**	J **A25**	H **A23**	G **A22**
K **A22**	G **A18**	P **A27**	I **A23**	M **A27**	L **A26**	M **A28**	N **A29**	I **A24**
0 **A26**	I **A20**	R **A29**	K **A25**	N **A28**	P **A29**	O **A30**	Q **A32**	K **A26**
Q **A28**	L **A23**	S **A30**	Q **A31**	O **A29**	R **A32**			L **A27**
U **A32**	N **A25**	T **A31**			U **B3**			P **A31**
V **B2**	a **B7**	W **B3**	a **B9**	S **B1**	X **B6**	R **B1**	T **B3**	U **B4**
Z **B6**	c **B9**	X **B4**	d **B12**	T **B2**	Z **B8**	S **B2**	V **B5**	X **B7**
f **B12**	d **B10**	Y **B5**	f **B14**	V **B4**	b **B10**	Y **B8**	W **B6**	c **B12**
g **B13**	e **B11**	b **B8**	g **B15**	W **B5**	c **B11**	Z **B9**	b **B11**	e **B14**
i **B15**	j **B16**	h **B14**	h **B16**	Y **B7**	f **B13**	a **B10**	d **B13**	
						g **B16**	f **B15**	
m **C3**	k **C1**	o **C5**	l **C4**	i **C1**	j **C2**	l **C5**	h **C1**	j **C3**
n **C4**	l **C1**	p **C6**	m **C5**	n **C6**	k **C3**	m **C6**	i **C2**	k **C4**

4.-6. KLASSE

SeSam4			SeSam5			SeSam6		
SeSam4A	**SeSam4B**	**SeSam4C**	**SeSam5A**	**SeSam5B**	**SeSam5C**	**SeSam6A**	**SeSam6B**	**SeSam6C**
MP4A	**MP4B**	**MP4C**	**MP5 A**	**MP5 B**	**MP5 C**	**MP5 A**	**MP5 B**	**MP6 C**
E **A5**	B **A2**	A **A1**	A **A1**	I **A9**	G **A7**	B **A2**	D **A4**	A **A1**
F **A6**	G **A7**	C **A3**	B **A2**	J **A10**	H **A8**	I **A9**	E **A5**	C **A3**
I **A9**	H **A8**	D **A4**	C **A3**	K **A11**	L **A12**	J **A10**	K **A11**	F **A6**
J **A10**	M **A13**	S **A19**	D **A4**	M **A13**	N **A14**	L **A12**	M **A13**	G **A7**
K **A11**	O **A15**	T **A20**	E **A5**	S **A19**	P **A16**	R **A18**	N **A14**	H **A8**
L **A12**	P **A16**	U **A21**	F **A6**	T **A20**	U **A21**	S **A19**	P **A16**	O **A15**
N **A14**	R **A18**	V **A22**	O **A15**	V **A22**	X **A24**	U **A21**	T **A20**	Q **A17**
Q **A17**	W **A23**	Z **A26**	Q **A17**	W **A23**	Z **A26**	W **A23**	X **A24**	V **A22**

b	**A28**	X	**A24**	a	**A27**	R	**A18**	Y	**A25**	a	**A27**	a	**A27**	Y	**A25**	Z	**A26**
c	**A29**	Y	**A25**	d	**A30**	b	**A28**	d	**A30**	e	**A31**	d	**A30**	c	**A29**	b	**A28**
e	**A31**	f	**A32**			c	**A29**	f	**A32**			f	**A32**			e	**A31**
k	**B5**	h	**B2**	g	**B1**	g	**B1**	i	**B3**	h	**B2**	j	**B4**	g	**B1**	k	**B5**
o	**B9**	l	**B6**	i	**B3**	m	**B7**	J	**B4**	k	**B5**	l	**B6**	h	**B2**	m	**B7**
q	**B11**	n	**B8**	j	**B4**	r	**B12**	n	**B8**	l	**B6**	q	**B11**	i	**B3**	p	**B10**
t	**B14**	p	**B10**	m	**B7**	s	**B13**	t	**B14**	o	**B9**	s	**B13**	n	**B8**	u	**B15**
u	**B15**	s	**B13**	r	**B12**	v	**B16**	u	**B15**	p	**B10**	t	**B14**	o	**B9**	v	**B16**
				v	**B16**					q	**B11**			r	**B12**		
x	**C2**	y	**C3**	w	**C1**	w	**C1**	x	**C2**	y	**C3**	z	**C4**	w	**C1**	y	**C3**
ø	**C6**	z	**C4**	æ	**C5**	z	**C4**	ø	**C6**	æ	**C5**	ø	**C6**	x	**C2**	æ	**C5**

Analyserne er foretaget på samme for hvert af de forskellige delprojekter. Detaljerne i disse analyser vil bliver beskrevet i forbindelse med SeSam1 projektet for første klasse, hvorefter beskrivelsen af de øvrige delprojekter opsummerer resultater uden yderligere diskussion af metoderne.

4.SeSam1

4.1 MatematikProfilen 1

I dette afsnit vil vi omtale MatematikProfilen for 1. klasse som MP1 og SeSams tre deltest som henholdsvis SeSam1A, SeSam1B og SeSam1C.

Fra 1. til 3. klasse måler MatematikProfilen to forskellige dimensioner. Regne-færdigheder og forståelse af matematiske fænomener, men SeSam projektet anvender kun opgaver fra den sidste del, med en maksimal MP score på 69 point.

SeSam projektet anvender test-scores og kriteriebaserede kategorier som mål for færdighed i stedet for WML estimater af Rasch modellen personparameter. På trods af det inkluder vi tabel 4.1 med såkaldte vægtede likelihood (WML) estimater af person parametrene i dette og de efterfølgende afsnit. Det er der to grunde til.

Den første grund er at diskussioner af sikkerheden af målinger ved hjælp af Rasch modeller altid fokuserer på bias og standard fejlen (SEM) af estimater af målinger på logit-skalaen. Hvis vi derfor ønsker at sammenligne sikkerheden i MatematikProfilen med sikkerheden i andre test er vi nødt til inkludere WML-estimaterne.

I forbindelse med adaptive test er målet som regel at komme ned på SEM værdier tæt på 0,30. Der kom DNT sjældent ned, men tabel 2.2 viser at SEM kommer ned under 0,25 for testscores fra 22 til 51, at den er under 0,30 fra 14 til 56 og at den kun er over 0,40 fra testscores mellem 0 og 8 eller mellem 65 og 69. MatematikProfilen giver særdeles sikre målinger. Blandt de elever, der deltog i Gyldendals afprøvning af MatematikProfilen for 1. klasse, var det kun 3.4 % der havde SEM større end 0.40.

Den anden grund til at vi inkluderer tabeller for WML estimater af den samlede matematikprofil er at tabellen sidste søjle giver et estimat af SEM værdien for den samlede test-score. Da det er denne test-score, som SeSam benytter er det nyttigt at have en fornemmelse af usikkerhed på disse målinger. Sammenlignet med SEM værdierne for WML estimatet ser det ud som om der er tale om et paradoks. Usikkerhed på den samlede test-score er størst, der hvor SEM på WML estimatet er mindst. Det er imidlertid ikke et paradoks. Det skyldes at den samlede testskala er begrænsebåde nedadtil og opadtil, mens WML estimatets skal i princippet går fra minus til plus uendelig.

Tabel 4.1 Vægtede likelihood (WML) estimater af personparameteren i MP1.

Score	Theta estimate	True score	Bias	SEM	Score SEM
2	-4.801	2.48	0.490	0.806	0.69
3	-3.657	3.46	0.065	0.747	1.18
4	-3.113	4.46	0.009	0.676	1.52
5	-2.748	5.45	-0.000	0.592	1.79
6	-2.474	6.46	-0.002	0.522	2.03
7	-2.256	7.46	-0.002	0.469	2.25
8	-2.073	8.46	-0.002	0.428	2.44

9	-1.917	9.46	-0.001	0.395	2.62
10	-1.782	10.46	-0.001	0.370	2.79
11	-1.661	11.45	-0.001	0.348	2.95
12	-1.553	12.44	-0.001	0.331	3.10
13	-1.454	13.43	-0.001	0.316	3.23
14	-1.364	14.41	-0.001	0.304	3.36
15	-1.280	15.39	-0.001	0.293	3.47
16	-1.201	16.37	-0.001	0.284	3.58
17	-1.127	17.34	-0.001	0.276	3.67
18	-1.056	18.32	-0.000	0.269	3.76
19	-0.989	19.30	-0.000	0.263	3.84
20	-0.924	20.27	-0.000	0.258	3.91
21	-0.861	21.24	0.000	0.253	3.97
22	-0.800	22.22	0.000	0.249	4.03
23	-0.741	23.20	0.000	0.246	4.08
24	-0.683	24.17	0.000	0.243	4.12
25	-0.626	25.16	0.000	0.241	4.16
26	-0.569	26.14	0.000	0.239	4.20
27	-0.514	27.13	0.000	0.237	4.23
28	-0.459	28.11	0.000	0.235	4.26
29	-0.405	29.10	0.000	0.233	4.28
30	-0.351	30.09	0.000	0.232	4.31
31	-0.298	31.09	0.000	0.231	4.33
32	-0.245	32.08	0.000	0.230	4.35
33	-0.192	33.08	-0.000	0.229	4.37
34	-0.140	34.07	-0.000	0.228	4.38
35	-0.089	35.07	-0.000	0.227	4.40
36	-0.038	36.06	-0.000	0.226	4.41
37	0.013	37.06	-0.001	0.225	4.43
38	0.064	38.05	-0.001	0.225	4.44
39	0.113	39.03	-0.001	0.225	4.45
40	0.163	40.01	-0.001	0.225	4.45

Tabellen fortsættes på næste side

Score	Theta estimate	True score	Bias	RMSE	Score SEM
41	0.213	40.99	-0.001	0.225	4.45
42	0.262	41.97	-0.001	0.225	4.45
43	0.311	42.94	-0.001	0.226	4.44
44	0.360	43.91	-0.000	0.227	4.42
45	0.410	44.87	-0.000	0.228	4.40
46	0.460	45.84	-0.000	0.230	4.37
47	0.511	46.80	0.000	0.233	4.33
48	0.563	47.76	0.000	0.236	4.28
49	0.616	48.72	0.001	0.239	4.22
50	0.671	49.68	0.001	0.244	4.15
51	0.728	50.65	0.001	0.249	4.07
52	0.788	51.62	0.001	0.255	3.98
53	0.851	52.59	0.001	0.261	3.88
54	0.917	53.56	0.001	0.269	3.77
55	0.988	54.55	0.001	0.278	3.66
56	1.065	55.53	0.001	0.289	3.53
57	1.148	56.52	0.001	0.301	3.39
58	1.238	57.52	0.001	0.315	3.25
59	1.337	58.52	0.001	0.331	3.10
60	1.447	59.52	0.001	0.350	2.94
61	1.570	60.52	0.001	0.373	2.77
62	1.709	61.53	0.001	0.400	2.60
63	1.871	62.54	0.001	0.434	2.41
64	2.060	63.55	0.001	0.476	2.21
65	2.289	64.56	0.001	0.531	1.99
66	2.575	65.56	-0.000	0.602	1.76
67	2.955	66.57	-0.010	0.688	1.49
68	3.519	67.55	-0.066	0.761	1.16
69	4.682	68.52	-0.495	0.816	0.68

4.2 Del-scores i MP1

Definitionen af SeSams deltest til 1. klase foregik på følgende måde.

Det første skridt gentog rutinemæssigt kontrollen af Gyldendals model for den samlede MP1 test og beregnede såkaldte vægtede likelihood estimater af personparameteren. Der blev ikke afsløret problemer med modellen.

Næste skridt beregnede tre delscores, MP1A, MP1B og MP1C ud fra Gyldendals data. Da der kun benyttes Gyldendals data til disse beregninger vil MP1 = MP1A + MP1B + MP1C.

Den log-lineære Rasch model for den samlede MP1 bruges derefter til at definerer tilsvarende modeller for hver af de tre del-scores. Disse modeller blive også rutinemæssigt afprøvet for at sikre, at der ikke dukkede modelfejl op, som var overset i den samlede model. Disse modeller blev også accepteret, hvorefter der blev beregnet estimater af personparametrene for hvert af de enkelte del-scores. De forskellige del-estimater måler færdigheden på den samme logit-skala som det samlede estimat, men de vil være mere usikre end det samlede estimat, og der er ikke nogen simpel lineær relation der definerer den samlede estimat ud fra de enkelte del-estimater.

Da det er hensigten at måle færdighederne på skalaen for den samlede MP score og at bruge MatematikProfilens kategoriseringer af skalaen til at fortolke testresultaterne med omregnes estimaterne af personparametrene for hvert af de tre del-scores til den forventede score på den samlede MP1.

Figur 4.1-4.3 resultaterne for MP1A, MP1B og MP1C. I disse figurer er de udvalgte opgaver markeret med variabel labels bestående af et enkelt bogstav som det program, der er brugt til beregningerne anvender. Disse labels anvendes til definitionen af de log-lineære Rasch modeller, der passer til deltestene.

Figur 4.1 indeholder oplysninger om deltest MP1A. Den øverste del definerer modellen med reference til ovennævnte label. Derefter følger definitionen af det samlede testresultat på MP1A efterfulgt af oplysningerne om, hvorledes del-skalaen fungerer og om omregningen til den samlede MP1 skala.

Den første søjle indeholder MP1A scoren. Der kan højst scores 23 point.

Søjle 2-4 viser WML estimaterne der måler færdighederne på MP1s logit skala. Målingerne vha. MP1A er uden nævneværdig bias, hvis der ikke er tale om ekstreme test-scores. SEM er naturligvis svagere end for den samlede MP1, men sikkerheden er sammenlignelig med det, som de tidligere nationale test kunne præstere.

Søjle 5-7 vise de såkaldte maksimum likelihood (ML) estimaterne. Disse resultater bruges til beregning af den samlede forventede test-score, men uden interesse i sig selv.

Søjle 8 viser den forventede score på de opgaver, der ikke er med i MP1A, hvorefter søjle 9 viser den forventede MP1 score som decimaltal, mens søjle 19 viser det samme tal afrundet uden decimaler. Vi bruger decimaltallet i forbindelse med statistiske analyser af resultaterne, men forventer at lærere, der har i sinde at bruge resultaterne på elev eller klasseniveau kun har brug for det afrundede tal.

Et enkelt eksempel kan illustrere hvorledes resultaterne hænger sammen.

Antag, at el elev har scoret 8 ud af 23 point på MP1A. Det svarer til et mål på MP1s logit-skala lig med -0.579 med SEM lig med 0,448. Ifølge tabel 2.1 og 2.2 vil en på dette niveau forvente en samlet score omkring 26 point på den samlede MP1, hvor ML estimatet vil være -0.577 (SEM = 0,244) mens WML vil være -0,569 (SEM = 0,239).

Den ottende søjle i Figur 2.2 viser, at den forventede score på de opgaver, der ikke er inkluderet i MP1A er lig med 16,7, således at den forventede samlede score er lig med 24,678 eller – afrundet – lig med 25.

```
Observed subscore:
A2+A3+A21+A22+A26+A28+A32+B2+B6+B12+B13+B15+C3+C4

Expected
           WML estimates           ML estimates
missing   total
Score   Theta    Bias     SEM   Theta    Bias     SEM
subscore   score
-----------------------------------------------------------
-----------
```

1	-4.015	0.515	0.700	-4.610	0.359	0.678	1.3	2.344	2
2	-2.737	0.068	0.831	-3.104	-0.256	0.987	2.5	4.530	5
3	-2.067	0.010	0.748	-2.274	-0.247	0.927	4.5	7.453	7
4	-1.611	-0.001	0.650	-1.759	-0.185	0.792	6.7	10.694	11
5	-1.271	-0.003	0.574	-1.384	-0.139	0.678	9.2	14.152	14
6	-1.002	-0.004	0.519	-1.089	-0.107	0.596	11.7	17.711	18
7	-0.776	-0.003	0.479	-0.845	-0.083	0.537	14.3	21.251	21
8	-0.579	-0.003	0.448	-0.633	-0.065	0.494	16.7	24.678	25
9	-0.403	-0.003	0.425	-0.445	-0.050	0.464	19.0	27.953	28
10	-0.241	-0.003	0.409	-0.273	-0.037	0.443	21.1	31.085	31
11	-0.090	-0.002	0.397	-0.112	-0.024	0.430	23.1	34.110	34
12	0.054	-0.002	0.391	0.042	-0.011	0.424	25.1	37.073	37
13	0.193	-0.001	0.390	0.192	0.002	0.426	27.0	40.016	40
14	0.330	-0.000	0.394	0.341	0.018	0.435	29.0	42.967	43
15	0.470	0.002	0.404	0.495	0.036	0.454	30.9	45.937	46
16	0.616	0.003	0.420	0.657	0.058	0.483	32.9	48.913	49
17	0.773	0.006	0.445	0.834	0.085	0.525	34.9	51.871	52
18	0.952	0.008	0.482	1.038	0.118	0.586	36.8	54.793	55
19	1.166	0.009	0.534	1.284	0.159	0.669	38.7	57.681	58
20	1.440	0.008	0.608	1.606	0.213	0.779	40.6	60.555	61
21	1.828	-0.003	0.711	2.069	0.276	0.907	42.4	63.435	63
22	2.462	-0.062	0.822	2.872	0.272	0.973	44.3	66.294	66
23	3.812	-0.531	0.721	4.394	-0.363	0.681	46.0	69.000	69

Figur 4.1 Log-lineær Rasch model for MP1A

Observed subscore:
A5+A6+A17+A18+A20+A23+A25+B7+B9+B10+B11+B16+C1+C2

Score	WML estimates Theta	Bias	SEM	ML estimates Theta	Bias	SEM	Expected missing subscore	total score	
1	-2.462	0.065	0.717	-2.876	-0.289	0.911	4.1	5.135	5
2	-1.949	0.005	0.644	-2.180	-0.285	0.872	5.9	7.942	8
3	-1.623	-0.007	0.564	-1.779	-0.217	0.758	7.5	10.541	11
4	-1.386	-0.009	0.502	-1.497	-0.159	0.653	9.0	12.993	13
5	-1.199	-0.007	0.458	-1.277	-0.115	0.572	10.3	15.348	15
6	-1.040	-0.005	0.427	-1.094	-0.080	0.512	11.7	17.651	18
7	-0.897	-0.002	0.405	-0.931	-0.053	0.470	12.9	19.939	20
8	-0.762	-0.000	0.392	-0.781	-0.031	0.441	14.2	22.248	22
9	-0.632	0.002	0.384	-0.637	-0.014	0.422	15.6	24.613	25
10	-0.500	0.003	0.381	-0.495	0.000	0.411	17.1	27.067	27
11	-0.365	0.003	0.382	-0.351	0.011	0.407	18.6	29.646	30
12	-0.224	0.003	0.387	-0.203	0.019	0.408	20.4	32.384	32
13	-0.074	0.002	0.395	-0.049	0.025	0.415	22.3	35.318	35
14	0.085	0.001	0.406	0.114	0.032	0.428	24.5	38.479	38
15	0.253	-0.000	0.420	0.286	0.039	0.449	26.9	41.877	42
16	0.432	-0.001	0.439	0.470	0.050	0.479	29.5	45.470	45
17	0.621	-0.000	0.463	0.670	0.068	0.522	32.1	49.147	49
18	0.826	0.001	0.496	0.892	0.095	0.585	34.7	52.749	53
19	1.056	0.004	0.542	1.148	0.137	0.670	37.2	56.168	56
20	1.328	0.005	0.603	1.463	0.198	0.778	39.4	59.390	59

21	1.681	-0.002	0.679	1.895	0.274	0.894	41.5	62.494	62
22	2.218	-0.057	0.751	2.625	0.283	0.935	43.6	65.619	66
23	3.413	-0.502	0.656	4.066	-0.341	0.645	46.0	69.000	69

Figur 4.2 Log-lineær Rasch model for MP1B del-score

Observed subscore:
A4+A19+A24+A27+A29+A30+A31+B3+B4+B5+B8+B14+C5+C6

	WML estimates			ML estimates					
Score	Theta	Bias	SEM	Theta	Bias	SEM	Expected missing subscore	total score	
1	-3.519	0.486	0.631	-4.196	0.333	0.634	1.6	2.640	3
2	-2.401	0.064	0.740	-2.786	-0.277	0.931	3.4	5.406	5
3	-1.860	0.007	0.684	-2.059	-0.261	0.897	5.6	8.634	9
4	-1.485	-0.002	0.612	-1.613	-0.185	0.779	7.9	11.908	12
5	-1.190	-0.001	0.553	-1.281	-0.129	0.668	10.3	15.311	15
6	-0.940	0.000	0.508	-1.008	-0.092	0.584	12.8	18.828	19
7	-0.719	0.001	0.474	-0.773	-0.069	0.524	15.4	22.385	22
8	-0.516	0.001	0.446	-0.562	-0.056	0.480	17.9	25.891	26
9	-0.328	-0.001	0.423	-0.371	-0.047	0.449	20.3	29.281	29
10	-0.153	-0.004	0.403	-0.196	-0.041	0.428	22.5	32.516	33
11	0.004	-0.006	0.389	-0.035	-0.035	0.416	24.6	35.580	36
12	0.144	-0.006	0.380	0.113	-0.025	0.414	26.5	38.475	38
13	0.271	-0.005	0.377	0.253	-0.010	0.419	28.2	41.221	41
14	0.388	-0.002	0.380	0.387	0.010	0.432	29.9	43.855	44
15	0.502	0.002	0.388	0.520	0.036	0.453	31.4	46.422	46
16	0.619	0.006	0.402	0.660	0.065	0.481	33.0	48.972	49
17	0.746	0.010	0.424	0.814	0.098	0.520	34.6	51.558	52
18	0.894	0.013	0.457	0.996	0.133	0.574	36.2	54.235	54
19	1.084	0.012	0.508	1.225	0.172	0.653	38.0	57.048	57
20	1.351	0.006	0.587	1.536	0.218	0.763	40.0	60.004	60

```
 21     1.751  -0.011   0.699   1.993   0.273   0.897   42.0
63.040  63
 22     2.379  -0.072   0.800   2.773   0.272   0.959   44.0
66.042  66
 23     3.633  -0.514   0.689   4.250  -0.351   0.663   46.0
69.000  69
--------------------------------------------------------------
-----------
```

Figur 4.3 Log-lineær Rasch model for MP1C del-score

4.3 SeSam1

Tabel 4.2 viser de forventede scores for MatematikProfilen for 1. klasse3.

En samlet SeSam1A score lig med 10 modsvarer en forventet delscore på MP1A lig med 9.23 og en samles score på MP1 lig med 28.67. En tilsvarende SeSam1B score modsvarer en forventet delscore på MP1B lig med 9.50 og en samles score på MP1 lig med 25.84. For SeSam1C er tallene lig med 12.76 og 40.564.

Tabel 4.2 Oversigt over forventede MP scores

Score på SeSam1 A-C	SeSam1A		SeSam1B		SeSam1C	
	E(MP1A)	E(MP1)	E(MP1B)	E(MP1)	E(MP1C)	E(MP1)
0	0	0	0	0	0	0
1	0.89	2.08	0.23	1.81	1.06	2.81
2	1.74	3.96	0.86	4.42	2.22	6.12
3	2.58	6.23	1.66	6.99	3.42	10.01
4	3.41	8.78	2.49	9.21	4.63	14.05
5	4.27	11.63	3.39	11.50	5.85	18.30
6	5.15	14.72	4.41	13.96	7.09	22.70
7	6.10	18.06	5.55	16.62	8.36	27.11
8	7.09	21.56	6.80	19.48	9.70	31.55
9	8.14	25.14	8.13	22.56	11.15	36.01
10	9.23	28.67	9.50	25.84	12.76	40.56
11	10.36	32.17	10.86	29.29	14.47	45.06
12	11.52	35.65	12.20	32.97	16.09	49.21
13	12.68	39.07	13.50	36.90	17.43	52.71
14	13.84	42.49	14.78	41.12	18.44	55.47
15	14.99	45.91	16.03	45.58	19.18	57.58
16	16.12	49.27	17.22	49.94	19.77	59.32
17	17.23	52.54	18.34	53.91	20.26	60.79
18	18.30	55.66	19.35	57.30	20.71	62.16
19	19.34	58.66	20.23	60.10	21.16	63.52
20	20.33	61.51	21.02	62.56	21.64	64.96
21	21.27	64.21	21.73	64.78	22.14	66.46
22	22.16	66.73	22.39	66.94	22.60	67.82
23	23	69	23	69	23	69

[3] De forventede scores er decimaltal, men brugere af testene vil formodentlig være tilfredse med heltal
[4] Forskellen på SeSam1C og MP1C skydes at visse opgaver (især C5 og C6) er vanskeligere for SeSams elever end for dem der deltog i afprøvningen af Matametikprofilen

4.4 De gentagne SeSam1 målinger

Dette afsnits analysen og sammenligning af gentagne målinger inkluderer kun elever, der har deltaget i alle tre test.

Tabel 4.3 Gennemsnitlige testresultater målt på den samlede MP skala

Skole	n	SeSam1A E(MP1)	SeSam1B E(MP1)	SESEAM1C E(MP1)	Parvis t-test A & B	Parvis t-test B & C
N	22	40.4	52.2	61.1	<0.0005	<0.0005
S	17	37.7	35.6	39.9	0.25	0.046
H	21	39.0	43.9	54.6	0.039	<0.0005
R	24	49.9	45.4	59.0	0.032	<0.0005
I alt	84	42.2	44.8	54.6	0.007	<0.0005

Der er samlet set en signifikant forskel på SeSam1A og SeSam1B og en højsignifikant forskel mellem SeSam1B og SeSam1C, men tendenserne er forskellige, når vi ser på de fire forskellige skoler der deltog i projektet. På skole S er den ingen forskel på A og B og kun en svagt signifikant forskel på B og C.

Det interessante spørgsmål er om ændringerne fra B til C er større end ændringerne fra A til B. Målt ud fra den forventede MP1 score ser det ud til At være tilfældet, men for at fortolke ændringerne i forhold til elevernes færdigheder er det nødvendigt at se på, de kriteriebasered tolkninger i Tabel 4.3.

Tabel 4.3 Procentvis fordeling af elever ifølge MatematikProfilens kriteriebaserede kategorier

Skole	SeSam1	Begyndt	I gang	Godt i gang	Længere end forventet	Meget længere end forventet
N	SeSam1A	9.1 %	68.2 %	13.6 %	9.1 %	
	SeSam1B		22.7 %	27.3 %	18.2 %	31.8 %
	SeSam1C		4.5 %	4.5 %	9.1 %	81.8 %
S	SeSam1A	29.4 %	41.2 %	23.5 %	5.9 %	
	SeSam1B	17.6 %	58.8 %	23.5 %		
	SeSam1C	11.8 %	29.4 %	58.8 %		
H	SeSam1A	14.3 %	61.9 %	19.0 %	4.8 %	
	SeSam1B		42.9 %	47.6 %		9.5 %
	SeSam1C		4.8 %	28.6 %	38.1 %	28.6 %
R	SeSam1A		29.2 %	33.3 %	16.7 %	20.8 %
	SeSam1B	8.3 %	37.5 %	37.5 %	8.3 %	
	SeSam1C		4.2 %	8.3 %	20.8 %	66.7 %
I alt	SeSam1A	11.9 %	50.0 %	22.6 %	9.5 %	6.0 %
	SeSam1B	6.0 %	39.3 %	34.5 %	7.1 %	13.1 %
	SeSam1C	2.4 %	9.5 %	22.6 %	17.9 %	47.6 %

5.SeSam2

5.1 MatematikProfilen 2

Resultaterne for SeSam2 præsenteres på samme måde som for SeSam1, men uden kommentarer.

Tabel 5.1 WML estimater af personparameteren i MP2.

Score	Theta estimate	True score	Bias	SEM	Score SEM
0	-4.874	0.48	0.492	0.810	0.69
1	-3.723	1.46	0.066	0.753	1.17
2	-3.170	2.44	0.010	0.682	1.51
3	-2.799	3.44	0.000	0.597	1.77
4	-2.517	4.45	-0.001	0.527	2.01
5	-2.293	5.45	-0.001	0.473	2.23
6	-2.106	6.46	-0.001	0.431	2.42
7	-1.947	7.47	-0.001	0.398	2.61
8	-1.809	8.47	-0.001	0.371	2.79
9	-1.688	9.48	-0.001	0.349	2.95
10	-1.579	10.47	-0.001	0.330	3.11
11	-1.481	11.47	-0.001	0.314	3.25
12	-1.391	12.46	-0.001	0.301	3.39
13	-1.309	13.44	-0.001	0.289	3.52
14	-1.232	14.43	-0.001	0.279	3.64
15	-1.160	15.41	-0.001	0.271	3.75
16	-1.093	16.38	-0.001	0.263	3.85
17	-1.029	17.35	-0.001	0.257	3.94
18	-0.968	18.32	-0.001	0.251	4.03
19	-0.909	19.29	-0.000	0.246	4.10
20	-0.853	20.26	-0.000	0.242	4.16
21	-0.798	21.23	-0.000	0.239	4.22
22	-0.744	22.20	0.000	0.236	4.27
23	-0.691	23.17	0.000	0.233	4.31
24	-0.639	24.14	0.000	0.231	4.34
25	-0.588	25.12	0.000	0.229	4.37
26	-0.537	26.09	0.000	0.228	4.40
27	-0.486	27.07	0.000	0.227	4.42
28	-0.436	28.06	0.000	0.226	4.43
29	-0.386	29.05	0.000	0.225	4.44
30	-0.336	30.03	0.000	0.224	4.45
31	-0.286	31.02	0.000	0.224	4.46
32	-0.236	32.01	0.000	0.224	4.46
33	-0.186	33.01	0.000	0.224	4.46
34	-0.137	34.00	-0.000	0.224	4.46
35	-0.087	34.99	-0.000	0.224	4.46
36	-0.037	35.98	-0.000	0.224	4.46
37	0.013	36.96	-0.000	0.225	4.45
38	0.062	37.95	-0.000	0.225	4.44
39	0.113	38.94	-0.000	0.226	4.43
40	0.163	39.92	-0.000	0.227	4.41

Tabellen fortsættes på næste side

Score	Theta estimate	True score	Bias	SEM	Score SEM
41	0.213	40.90	-0.000	0.228	4.39
42	0.265	41.88	-0.000	0.229	4.37
43	0.316	42.86	-0.000	0.231	4.34
44	0.368	43.84	-0.000	0.233	4.31
45	0.422	44.81	-0.000	0.235	4.27
46	0.476	45.79	0.000	0.238	4.22
47	0.531	46.76	0.000	0.241	4.17
48	0.588	47.74	0.000	0.245	4.11
49	0.647	48.71	0.000	0.249	4.04
50	0.707	49.69	0.000	0.254	3.97
51	0.771	50.67	0.001	0.260	3.89
52	0.837	51.64	0.001	0.266	3.80
53	0.906	52.62	0.001	0.273	3.71
54	0.980	53.61	0.001	0.281	3.60
55	1.058	54.59	0.001	0.291	3.49
56	1.141	55.58	0.001	0.302	3.37
57	1.232	56.57	0.001	0.314	3.24
58	1.330	57.56	0.001	0.329	3.11
59	1.438	58.56	0.001	0.346	2.96
60	1.558	59.55	0.001	0.367	2.81
61	1.692	60.55	0.001	0.392	2.65
62	1.845	61.55	0.001	0.423	2.47
63	2.021	62.54	0.001	0.462	2.28
64	2.231	63.54	0.002	0.514	2.07
65	2.491	64.53	0.001	0.579	1.84
66	2.835	65.52	-0.009	0.657	1.56
67	3.352	66.51	-0.066	0.723	1.21
68	4.452	67.50	-0.484	0.788	0.71

Subscore = ABGIKQadfghlm

The GLLRM : Aa DY EF Qg ST Sn jk
Subscore model: Aa Qg

Observed subscore:
A15+A16+A21+A23+A25+A31+B9+B12+B14+B15+B16+C4+C5

	WML estimates			ML estimates			Expected		
Score	Theta	Bias	SEM	Theta	Bias	SEM	missing	total subscore	score
0	-3.938	0.513	0.696	-4.536	0.357	0.674	0.0	0.000	0
1	-2.668	0.068	0.825	-3.039	-0.259	0.981	1.8	2.761	3
2	-2.010	0.010	0.742	-2.218	-0.250	0.923	3.8	5.838	6
3	-1.559	-0.001	0.642	-1.711	-0.188	0.787	6.3	9.277	9
4	-1.226	-0.004	0.565	-1.344	-0.143	0.673	9.0	13.011	13
5	-0.967	-0.005	0.510	-1.058	-0.111	0.591	11.9	16.900	17
6	-0.753	-0.004	0.469	-0.823	-0.086	0.532	14.8	20.770	21
7	-0.568	-0.004	0.440	-0.621	-0.066	0.490	17.5	24.476	24
8	-0.401	-0.003	0.417	-0.441	-0.049	0.460	20.0	27.955	28
9	-0.247	-0.003	0.402	-0.276	-0.035	0.439	22.2	31.218	31
10	-0.103	-0.002	0.392	-0.121	-0.021	0.426	24.3	34.307	34
11	0.036	-0.002	0.387	0.028	-0.007	0.422	26.3	37.269	37
12	0.171	-0.001	0.388	0.175	0.007	0.425	28.1	40.147	40
13	0.306	0.001	0.394	0.322	0.023	0.435	30.0	42.974	43
14	0.444	0.003	0.405	0.475	0.042	0.453	31.8	45.780	46
15	0.591	0.005	0.422	0.639	0.062	0.480	33.6	48.590	49
16	0.753	0.006	0.447	0.821	0.085	0.518	35.4	51.422	51

```
 17      0.941   0.007   0.483   1.033   0.112   0.573   37.3
54.288  54
 18      1.173   0.005   0.533   1.291   0.145   0.654   39.2
57.177  57
 19      1.474  -0.000   0.604   1.622   0.192   0.767   41.0
60.048  60
 20      1.876  -0.013   0.691   2.077   0.257   0.895   42.8
62.826  63
 21      2.438  -0.070   0.749   2.811   0.273   0.933   44.5
65.462  65
 22      3.541  -0.481   0.629   4.213  -0.331   0.632   46.0
68.000  68
---------------------------------------------------------------
-----------
```

Figur 5.1 Log-lineær Rasch model for MP2A

MP2B

```
+-------------------------------------------------------+
|                                                       |
| Estimates of Person parameters based on subscore      |
|                                                       |
+-------------------------------------------------------+
```

Subscore = CDJMNOSTVWYin

The GLLRM : Aa DY EF Qg ST Sn jk
Subscore model: DY ST Sn

Observed subscore:
A17+A18+A24+A27+A28+A29+B1+B2+B4+B5+B7+C1+C6

	WML estimates			ML estimates			Expected		
Score	Theta	Bias	SEM	Theta	Bias	SEM	missing	subscore	total score
0	-3.511	0.499	0.637	-4.199	0.331	0.628	0.0	0.000	0
1	-2.354	0.059	0.707	-2.793	-0.299	0.903	2.5	3.460	3
2	-1.871	0.002	0.625	-2.115	-0.301	0.863	4.4	6.413	6
3	-1.575	-0.010	0.545	-1.738	-0.232	0.753	6.1	9.050	9
4	-1.363	-0.011	0.483	-1.478	-0.172	0.651	7.5	11.494	11
5	-1.196	-0.010	0.438	-1.279	-0.125	0.569	8.8	13.817	14
6	-1.056	-0.007	0.407	-1.114	-0.088	0.508	10.1	16.073	16
7	-0.932	-0.004	0.386	-0.969	-0.058	0.464	11.3	18.301	18
8	-0.818	-0.001	0.373	-0.836	-0.032	0.435	12.5	20.539	21
9	-0.707	0.001	0.366	-0.709	-0.011	0.416	13.8	22.827	23
10	-0.595	0.004	0.365	-0.583	0.008	0.406	15.2	25.204	25
11	-0.478	0.005	0.368	-0.454	0.023	0.404	16.7	27.716	28
12	-0.353	0.005	0.376	-0.317	0.035	0.408	18.4	30.413	30

13	-0.214	0.005	0.389	-0.169	0.044	0.420	20.3	33.345	33
14	-0.058	0.003	0.408	-0.008	0.053	0.441	22.6	36.558	37
15	0.116	0.001	0.431	0.171	0.062	0.472	25.1	40.075	40
16	0.309	-0.000	0.461	0.371	0.076	0.516	27.9	43.881	44
17	0.525	-0.000	0.499	0.598	0.098	0.579	30.9	47.906	48
18	0.769	0.001	0.547	0.863	0.134	0.666	34.0	52.026	52
19	1.058	0.002	0.610	1.190	0.191	0.777	37.1	56.114	56
20	1.426	-0.006	0.684	1.631	0.265	0.895	40.1	60.112	60
21	1.968	-0.062	0.746	2.361	0.279	0.932	43.1	64.063	64
22	3.114	-0.491	0.640	3.779	-0.335	0.637	46.0	68.000	68

Figur 5.2 Log-lineær Rach model for MP2B

MP2C

```
+---------------------------------------------------+
|                                                   |
| Estimates of Person parameters based on subscore |
|                                                   |
+---------------------------------------------------+
```

Subscore = EFHLPRUXZbcejk

The GLLRM   : Aa DY EF Qg ST Sn jk
Subscore model: EF jk

Observed subscore:
A19+A20+A22+A26+A30+A32+B3+B6+B8+B10+B11+B13+C2+C3

	WML estimates			ML estimates			Expected	
Score	Theta	Bias	SEM	Theta	Bias	SEM	missing subscore	total score
0	-3.859	0.509	0.678	-4.482	0.349	0.659	0.0	0.000 0
1	-2.621	0.062	0.787	-3.014	-0.273	0.956	1.8	2.826 3
2	-2.026	0.005	0.706	-2.242	-0.265	0.906	3.7	5.711 6
3	-1.635	-0.004	0.618	-1.780	-0.196	0.781	5.7	8.702 9
4	-1.342	-0.004	0.549	-1.446	-0.142	0.669	7.8	11.837 12
5	-1.102	-0.003	0.499	-1.181	-0.104	0.585	10.1	15.118 15
6	-0.896	-0.002	0.461	-0.957	-0.079	0.523	12.5	18.499 18
7	-0.712	-0.001	0.434	-0.761	-0.060	0.479	14.9	21.891 22
8	-0.544	-0.001	0.413	-0.583	-0.047	0.448	17.2	25.203 25
9	-0.389	-0.001	0.397	-0.420	-0.036	0.426	19.4	28.380 28
10	-0.242	-0.002	0.386	-0.266	-0.027	0.411	21.4	31.413 31
11	-0.103	-0.002	0.378	-0.121	-0.019	0.402	23.3	34.314 34
12	0.030	-0.002	0.373	0.020	-0.010	0.397	25.1	37.106 37

13	0.160	-0.001	0.372	0.157	-0.000	0.398	26.8	39.809	40
14	0.288	-0.001	0.374	0.294	0.011	0.404	28.4	42.439	42
15	0.416	0.000	0.379	0.432	0.023	0.416	30.0	45.011	45
16	0.548	0.001	0.390	0.576	0.039	0.436	31.5	47.534	48
17	0.687	0.003	0.406	0.729	0.058	0.465	33.0	50.021	50
18	0.837	0.004	0.429	0.896	0.081	0.508	34.5	52.482	52
19	1.007	0.006	0.462	1.086	0.112	0.569	35.9	54.932	55
20	1.206	0.006	0.508	1.313	0.154	0.653	37.4	57.391	57
21	1.452	0.005	0.570	1.601	0.211	0.760	38.9	59.886	60
22	1.782	-0.006	0.646	2.005	0.282	0.874	40.5	62.459	62
23	2.288	-0.064	0.713	2.697	0.289	0.909	42.2	65.160	65
24	3.392	-0.485	0.622	4.086	-0.327	0.623	44.0	68.000	68

Figur 5.3 Log-lineær Rasch model for MP2C

5.3 SeSam2

Omregning af observerede SeSam2 scores til forventede MP2 scores foregik på samme måde som med SeSam1.

Tabel 5.2 Oversigt over forventede MP scores

Score på SeSam2 A-C	SeSam2A		SeSam2B		SeSam2C	
	E(MP2A)	E(MP2)	E(MP2B)	E(MP2)	E(MP2C)	E(MP2)
0	0	0	0	0	0	0
1	0.39	1.08	0.15	0.52	0.88	2.49
2	0.82	2.64	0.62	2.14	1.73	4.93
3	1.27	3.59	1.34	4.64	2.60	7.51
4	1.72	4.98	2.19	6.91	3.52	10.33
5	2.20	6.53	3.20	9.54	4.47	13.38
6	2.70	8.24	4.41	12.45	5.45	16.64
7	3.24	10.17	5.81	15.64	6.47	20.09
8	3.81	12.30	7.37	19.13	7.52	23.61
9	4.44	14.72	8.98	22.78	8.59	27.08
10	5.11	17.33	10.51	26.48	9.68	30.44
11	5.85	20.19	11.91	30.17	10.79	33.71
12	6.67	23.25	13.16	33.86	11.90	36.83
13	7.58	26.49	14.28	37.54	13.01	39.84
14	8.59	29.88	15.29	41.18	14.12	42.75
15	9.74	33.50	16.24	44.85	15.22	45.57
16	11.05	37.41	17.12	48.40	16.31	48.31
17	12.55	41.70	17.96	51.86	17.38	50.96
18	14.24	46.45	18.77	55.17	18.43	53.54
19	16.08	51.65	19.56	58.35	19.46	56.06
20	17.98	57.12	20.35	61.49	20.46	58.54
21	19.90	62.55	21.15	64.65	21.43	60.99
22	22	68	22	68	22.35	63.40
23					23.21	65.76
4					24	69

Påfaldende dårlig tilpasning mellem scoren på SeSam2A og den forventede MP2A score. Resultaterne antyder opgaverne i SeSam2A var væsentlig lettere for de elever, der deltog i det første SeSam2 test end de var for eleverne der besvarede hele MP2.

5.4 De gentagne SeSam2 målinger

Analysen af de gentagne inkluderer kun elever, der har deltaget i alle tre test.

Tabel 5.3 Gennemsnitlige testresultater målt på den samlede MP skala

Skole	n	SeSam2A E(MP2)	SeSam2B E(MP2)	SESEAM2C E(MP2)	Parvis t-test A & B	Parvis t-test B & C
N	25	38.8	49.3	55.4	<0.0005	0.001
S	24	21.6	30.8	47.2	<0.0005	<0.0005
H	44	42.5	45.1	49.4	.074	.026
R	42	39.2	38.6	48.0	.607	<0.0005
I alt	135	37.1	41.3	49.7	<0.0005	<0.0005

Tendensen er overordnet som forventet, men det er påfaldende, at der ikke er forskel på resultaterne fra SeSam2A og SeSam2B på H og R skolerne. Det antyder, at eleverne på disse to skoler har haft påfaldende let ved opgaverne i SeSam2A.

Tabel 5.3 Procentvis fordeling af elever ifølge Matematikprofilens kriteriebaserede kategorier

Skole	SeSam2	Begyndt	I gang	Godt i gang	Længere end forventet	Meget længere end forventet
N	SeSam2A	8.0	48.0	36.0	8.0	
	SeSam2B	4.0	4.0	20.0	64.0	8.0
	SeSam2C			12.0	60.0	28.0
S	SeSam2A	79.2	12.5	8.3		
	SeSam2B	50.0	20.8	8.3	20.8	
	SeSam2C		20.8	29.2	37.5	12.5
H	SeSam2A	9.1	29.5	31.8	15.9	13.6
	SeSam2B	18.2	9.1	13.6	43.2	15.9
	SeSam2C	4.5	20.5	11.4	31.8	31.8
R	SeSam2A	19.0	33.3	28.6	7.1	11.9
	SeSam2B	19.0	33.3	19.0	23.8	4.8
	SeSam2C	4.8	9.5	23.8	52.4	9.5
I alt	SeSam2A	24.4	31.1	27.4	8.9	8.1
	SeSam2B	21.5	17.8	15.6	37.0	8.1
	SeSam2C	3.0	13.3	18.5	44.5	20.7

6.SeSam3

6.1 Matematikprofilen 3

Tabel 6.1 WML estimater af personparameteren i MP2.

Score	Theta estimate	True score	Bias	SEM	Score SEM
0	-5.423	0.46	0.498	0.831	0.66
1	-4.224	1.40	0.064	0.791	1.12
2	-3.621	2.37	0.009	0.722	1.41
3	-3.201	3.35	0.001	0.636	1.65
4	-2.874	4.35	0.000	0.565	1.85
5	-2.606	5.35	-0.000	0.511	2.03
6	-2.380	6.36	-0.000	0.469	2.20
7	-2.185	7.36	-0.001	0.435	2.35
8	-2.015	8.37	-0.001	0.408	2.51
9	-1.864	9.37	-0.001	0.385	2.65
10	-1.729	10.37	-0.001	0.366	2.78
11	-1.607	11.36	-0.001	0.349	2.91
12	-1.495	12.35	-0.001	0.335	3.03
13	-1.391	13.34	-0.001	0.323	3.14
14	-1.294	14.32	-0.001	0.312	3.24
15	-1.204	15.30	-0.001	0.303	3.33
16	-1.118	16.28	-0.000	0.295	3.42
17	-1.036	17.26	-0.000	0.288	3.50
18	-0.957	18.24	-0.000	0.283	3.56
19	-0.881	19.22	-0.000	0.277	3.63
20	-0.808	20.20	0.000	0.273	3.68
21	-0.737	21.18	0.000	0.269	3.73
22	-0.667	22.16	0.000	0.265	3.78
23	-0.599	23.15	0.000	0.262	3.82
24	-0.532	24.13	0.000	0.260	3.85
25	-0.466	25.12	0.000	0.258	3.89
26	-0.401	26.11	0.000	0.256	3.91
27	-0.337	27.10	0.000	0.254	3.94
28	-0.273	28.09	0.000	0.252	3.96
29	-0.210	29.08	0.000	0.251	3.98
30	-0.148	30.07	0.000	0.250	4.00
31	-0.087	31.06	-0.000	0.249	4.02
32	-0.025	32.05	-0.000	0.248	4.03
33	0.036	33.05	-0.000	0.247	4.04
34	0.096	34.04	-0.000	0.247	4.05
35	0.156	35.03	-0.000	0.246	4.06
36	0.216	36.02	-0.000	0.246	4.07
37	0.276	37.00	-0.000	0.246	4.07
38	0.335	37.99	-0.000	0.246	4.07
39	0.395	38.98	-0.000	0.246	4.07
40	0.454	39.96	0.000	0.247	4.06

Tabellen fortsættes på næste side

Score	Theta estimate	True score	Bias	SEM	Score SEM
41	0.514	40.95	0.000	0.247	4.05
42	0.575	41.93	0.000	0.248	4.03
43	0.636	42.92	0.000	0.249	4.01
44	0.697	43.91	0.000	0.251	3.99
45	0.760	44.90	0.000	0.252	3.97
46	0.823	45.89	0.000	0.253	3.94
47	0.888	46.88	-0.000	0.255	3.91
48	0.954	47.88	-0.000	0.257	3.88
49	1.020	48.88	-0.001	0.259	3.85
50	1.088	49.87	-0.001	0.261	3.82
51	1.157	50.87	-0.001	0.264	3.79
52	1.226	51.85	-0.002	0.267	3.75
53	1.296	52.83	-0.002	0.271	3.71
54	1.367	53.80	-0.001	0.276	3.66
55	1.440	54.75	-0.001	0.283	3.60
56	1.514	55.69	0.000	0.292	3.53
57	1.591	56.62	0.001	0.302	3.44
58	1.672	57.55	0.003	0.316	3.32
59	1.760	58.48	0.004	0.333	3.18
60	1.857	59.42	0.005	0.355	3.02
61	1.969	60.37	0.006	0.383	2.82
62	2.102	61.34	0.007	0.421	2.58
63	2.269	62.34	0.006	0.472	2.31
64	2.491	63.37	0.003	0.540	2.01
65	2.802	64.42	-0.010	0.624	1.66
66	3.295	65.46	-0.069	0.701	1.26
67	4.381	66.49	-0.483	0.781	0.72

MP3A

Observed subscore:
A19+A20+A25+A28+A30+B1+B2+B8+B9+B10+B16+C5+C6

	WML estimates			ML estimates				Expected missing subscore	total score
Score	Theta	Bias	SEM	Theta	Bias	SEM			
0	-4.086	0.554	0.766	-4.623	0.378	0.704	0.0	0.000	0
1	-2.608	0.056	0.847	-3.044	-0.275	0.988	2.8	3.801	4
2	-1.967	-0.003	0.716	-2.219	-0.286	0.914	5.2	7.180	7
3	-1.597	-0.011	0.615	-1.757	-0.219	0.790	7.2	10.153	10
4	-1.332	-0.009	0.544	-1.438	-0.157	0.682	8.9	12.876	13
5	-1.117	-0.005	0.494	-1.189	-0.109	0.597	10.5	15.463	15
6	-0.929	-0.001	0.459	-0.977	-0.074	0.532	12.0	17.982	18
7	-0.755	0.001	0.435	-0.788	-0.050	0.485	13.5	20.476	20
8	-0.588	0.002	0.417	-0.611	-0.034	0.452	15.0	22.970	23
9	-0.426	0.002	0.405	-0.442	-0.023	0.430	16.5	25.474	25
10	-0.266	0.001	0.397	-0.280	-0.017	0.415	18.0	27.987	28
11	-0.109	-0.001	0.391	-0.121	-0.012	0.408	19.5	30.499	30
12	0.043	-0.002	0.388	0.033	-0.006	0.407	21.0	33.002	33
13	0.189	-0.003	0.388	0.184	0.001	0.412	22.5	35.492	35
14	0.331	-0.003	0.391	0.334	0.012	0.423	24.0	37.976	38
15	0.471	-0.001	0.398	0.486	0.026	0.441	25.5	40.473	40
16	0.612	0.001	0.411	0.642	0.046	0.470	27.0	43.019	43
17	0.761	0.003	0.431	0.809	0.073	0.513	28.7	45.668	46
18	0.925	0.006	0.460	0.995	0.107	0.573	30.5	48.507	49

```
 19      1.112   0.008   0.501   1.214   0.154   0.655   32.7
51.679   52
 20      1.342   0.007   0.558   1.489   0.215   0.757   35.4
55.382   55
 21      1.651  -0.004   0.628   1.876   0.287   0.865   38.6
59.587   60
 22      2.131  -0.062   0.692   2.547   0.294   0.895   41.6
63.590   64
 23      3.207  -0.479   0.608   3.916  -0.321   0.613   44.0
67.000   67
----------------------------------------------------------------
          ----------
```

Figur 6.1 Log-lineær Rasch model for MP3A

Observed subscore:
A18+A21+A23+A29+A32+B3+B5+B6+B11+B13+B15+C1+C2

Score	WML estimates Theta	WML Bias	WML SEM	ML estimates Theta	ML Bias	ML SEM	Expected missing subscore	Expected total score
0	-4.504	0.510	0.697	-5.091	0.359	0.679	0.0	0.000	0
1	-3.234	0.066	0.840	-3.588	-0.250	0.995	1.4	2.438	2
2	-2.557	0.011	0.769	-2.743	-0.232	0.940	2.8	4.814	5
3	-2.067	0.001	0.675	-2.199	-0.167	0.804	4.3	7.288	7
4	-1.685	-0.002	0.601	-1.789	-0.124	0.690	5.9	9.912	10
5	-1.376	-0.003	0.547	-1.460	-0.097	0.611	7.7	12.672	13
6	-1.117	-0.003	0.508	-1.184	-0.076	0.556	9.5	15.523	16
7	-0.893	-0.003	0.479	-0.943	-0.059	0.518	11.4	18.414	18
8	-0.690	-0.002	0.458	-0.728	-0.044	0.490	13.3	21.306	21
9	-0.502	-0.001	0.442	-0.528	-0.032	0.469	15.2	24.184	24
10	-0.323	0.000	0.431	-0.340	-0.021	0.454	17.0	27.048	27
11	-0.149	0.000	0.424	-0.158	-0.012	0.444	18.9	29.911	30
12	0.023	0.000	0.420	0.020	-0.004	0.439	20.8	32.793	33
13	0.195	-0.000	0.420	0.197	0.003	0.439	22.7	35.705	36
14	0.367	-0.001	0.422	0.375	0.011	0.444	24.7	38.650	39
15	0.542	-0.001	0.428	0.556	0.020	0.457	26.6	41.623	42
16	0.719	-0.001	0.438	0.742	0.034	0.481	28.6	44.615	45
17	0.902	-0.001	0.455	0.938	0.054	0.519	30.6	47.641	48
18	1.096	0.001	0.481	1.149	0.085	0.578	32.8	50.755	51
19	1.307	0.003	0.517	1.386	0.131	0.660	35.1	54.051	54

```
 20      1.549   0.004   0.564   1.672   0.197   0.763   37.6
57.553  58
 21      1.853  -0.006   0.617   2.054   0.278   0.863   40.0
61.010  61
 22      2.295  -0.063   0.655   2.695   0.294   0.876   42.1
64.102  64
 23      3.266  -0.455   0.565   4.016  -0.308   0.592   44.0
67.000  67
---------------------------------------------------------------
                          -----------
```

Figur 6.2 Log-lineær Rach model for MP3B

Observed subscore:
A9+A17+A22+A24+A26+A27+A31+B4+B7+B12+B14+C3+C4

Score	WML estimates Theta	WML estimates Bias	WML estimates SEM	ML estimates Theta	ML estimates Bias	ML estimates SEM	Expected missing subscore	total score	
0	-4.459	0.554	0.788	-4.951	0.396	0.735	0.0	0.000	0
1	-2.942	0.061	0.924	-3.312	-0.241	1.044	2.1	3.063	3
2	-2.154	0.002	0.805	-2.368	-0.242	0.964	4.4	6.414	6
3	-1.646	-0.006	0.697	-1.788	-0.178	0.829	6.9	9.923	10
4	-1.267	-0.005	0.622	-1.363	-0.124	0.719	9.6	13.615	14
5	-0.953	-0.002	0.571	-1.019	-0.085	0.639	12.5	17.469	17
6	-0.676	-0.000	0.537	-0.720	-0.058	0.584	15.4	21.413	21
7	-0.419	0.002	0.516	-0.449	-0.039	0.546	18.4	25.380	25
8	-0.174	0.003	0.502	-0.194	-0.027	0.520	21.3	29.337	29
9	0.064	0.004	0.492	0.050	-0.020	0.500	24.3	33.278	33
10	0.299	0.005	0.483	0.287	-0.018	0.483	27.2	37.187	37
11	0.535	0.003	0.470	0.518	-0.021	0.468	30.0	41.000	41
12	0.774	-0.003	0.453	0.741	-0.026	0.458	32.6	44.594	45
13	0.997	-0.010	0.438	0.950	-0.027	0.462	34.8	47.835	48
14	1.184	-0.012	0.434	1.143	-0.014	0.490	36.7	50.670	51
15	1.338	-0.008	0.441	1.320	0.020	0.543	38.2	53.152	53
16	1.475	-0.001	0.457	1.490	0.076	0.612	39.4	55.389	55
17	1.610	0.008	0.478	1.667	0.153	0.688	40.5	57.493	57
18	1.755	0.014	0.504	1.874	0.243	0.763	41.6	59.570	60
19	1.937	0.010	0.535	2.165	0.332	0.826	42.7	61.746	62

20	2.234	-0.034	0.576	2.729	0.331	0.837	44.2	64.206	64
21	3.257	-0.475	0.559	4.057	-0.309	0.588	46.0	67.000	67

Figur 6.3 Log-lineær Rasch model for MP3C

6.3 SeSam3

Tabel 3.2 Oversigt over forventede MP scores

Score på SeSam3 A-C	SeSam3A		SeSam3B		SeSam3C	
	E(MP3A)	E(MP3)	E(MP3B)	E(MP3)	E(MP3C)	E(MP3)
0	0	0	0	0	0	0
1	1.28	4.75	1.36	3.29	0.88	2.70
2	2.10	7.48	2.56	6.20	1.86	5.95
3	3.21	10.72	3.71	9.15	2.91	9.61
4	4.32	13.70	4.84	12.23	3.98	13.54
5	5.43	16.55	6.00	15.41	5.04	17.63
6	6.52	19.28	7.04	18.53	6.03	21.53
7	7.57	21.90	8.08	21.54	6.97	25.18
8	8.58	24.42	9.04	24.30	7.79	28.51
9	9.55	26.86	9.94	26.88	8.56	31.54
10	10.47	29.17	10.78	29.28	9.28	34.37
11	11.36	31.40	11.58	31.58	9.97	37.07
12	12.22	33.55	12.35	33.81	10.65	39.66
13	13.07	35.67	13.12	36.06	11.35	42.26
14	13.91	37.75	13.90	38.36	12.10	44.92
15	14.76	39.87	14.71	40.76	12.93	47.61
16	15.63	42.08	15.57	43.33	13.90	50.39
17	16.53	44.42	16.49	46.10	15.08	53.33
18	17.46	46.97	17.48	49.13	16.57	56.59
19	18.44	49.03	18.55	52.57	18.26	60.14
20	19.48	53.46	19.72	56.57	19.82	63.76
21	20.57	57.78	20.94	60.80	21	67
22	21.72	62.47	22.11	64.42		
23	23	67	23	67		

Rimelig god overensstemmelse mellem SeSam3s deltests scores og MP3s delskalaer.

6.4 De gentagne SeSam3 målinger

Analysen af de gentagne inkluderer kun elever, der har deltaget i alle tre test.

Tabel 6.3 Gennemsnitlige testresultater målt på den samlede MP skala

Skole	n	SeSam3A E(MP3)	SeSam3B E(MP3)	SESEAM3C E(MP3)	Parvis t-test A & B	Parvis t-test B & C
N	26	53.6	35.8	48.0	*	<0.0005
S	22	30.8	34.8	33.2	0.055	0.294
H	21	45.4	39.3	45.9	0.993	0.005
R	-	-	-	-	-	-
I alt						

Her er der så tydeligt problemer med SeSam3A på skole N, at det ikke giver mening at beregne et samlet resultat. Og det er ikke muligt at se nogen udvikling i resultaterne for Skole S.

Tabel 5.3 Procentvis fordeling af elever ifølge Matematikprofilens kriteriebaserede kategorier

Skole	SeSam3	Begyndt	I gang	Godt i gang	Længere end forventet	Meget længere end forventet
N	SeSam3A	*	*	*	*	*
	SeSam3B		57.7	42.3		
	SeSam3C		11.5	34.6	11.5	42.3
S	SeSam3A	40.9	40.9	18.2		
	SeSam3B	27.3	36.4	22.7	4.5	9.1
	SeSam3C	27.3	27.3	36.4	9.1	
H	SeSam3A		14.3	42.9	19.0	23.8
	SeSam3B		47.6	33.3	143	4.8
	SeSam3C		4.8	38.1	47.6	9.5
R	SeSam3A					
	SeSam3B					
	SeSam3C					
I alt	SeSam3A					
	SeSam3B					
	SeSam3C					

7.SeSam4

7.1 Matematikprofilen 4

Tabel 7.1 WML estimater af personparameteren i MP4.

Score	Theta estimate	True score	Bias	SEM	Score SEM
0	-5.266	0.49	0.487	0.798	0.70
1	-4.144	1.47	0.066	0.738	1.19
2	-3.608	2.45	0.009	0.672	1.52
3	-3.247	3.43	-0.001	0.594	1.78
4	-2.972	4.42	-0.001	0.528	2.00
5	-2.747	5.41	-0.001	0.477	2.20
6	-2.557	6.40	-0.001	0.437	2.37
7	-2.391	7.40	-0.000	0.407	2.53
8	-2.245	8.39	-0.000	0.382	2.68
9	-2.114	9.39	-0.000	0.361	2.83
10	-1.994	10.39	-0.000	0.344	2.96
11	-1.885	11.39	-0.000	0.329	3.09
12	-1.785	12.38	-0.000	0.316	3.21
13	-1.692	13.38	-0.000	0.305	3.33
14	-1.604	14.37	-0.000	0.295	3.44
15	-1.523	15.37	-0.000	0.286	3.54
16	-1.446	16.36	-0.000	0.277	3.64
17	-1.373	17.36	-0.000	0.270	3.74
18	-1.304	18.35	-0.000	0.263	3.83
19	-1.238	19.34	-0.000	0.257	3.92
20	-1.175	20.33	-0.000	0.252	4.01
21	-1.114	21.32	-0.000	0.246	4.09
22	-1.056	22.32	-0.000	0.242	4.16
23	-1.000	23.31	-0.000	0.237	4.24
24	-0.945	24.30	-0.000	0.233	4.31
25	-0.893	25.29	-0.000	0.230	4.37
26	-0.842	26.28	-0.000	0.226	4.44
27	-0.792	27.27	-0.000	0.223	4.50
28	-0.744	28.26	-0.000	0.220	4.55
29	-0.697	29.25	-0.000	0.218	4.61
30	-0.651	30.23	-0.000	0.215	4.66
31	-0.606	31.22	-0.000	0.213	4.71
32	-0.562	32.21	-0.000	0.211	4.75
33	-0.518	33.19	-0.000	0.209	4.79
34	-0.476	34.18	-0.000	0.207	4.83
35	-0.434	35.16	-0.000	0.206	4.86
36	-0.393	36.14	-0.000	0.205	4.90
37	-0.352	37.13	-0.000	0.203	4.92
38	-0.312	38.11	-0.000	0.202	4.95
39	-0.272	39.09	-0.000	0.202	4.96
40	-0.232	40.07	-0.000	0.201	4.98

Tabellen fortsættes på næste side

Score	Theta estimate	True score	Bias	SEM	Score SEM
41	-0.193	41.05	-0.000	0.200	4.99
42	-0.153	42.03	-0.000	0.200	5.00
43	-0.114	43.01	-0.000	0.200	5.00
44	-0.075	43.99	0.000	0.200	5.00
45	-0.036	44.97	0.000	0.200	5.00
46	0.003	45.95	0.000	0.200	4.99
47	0.043	46.93	0.000	0.201	4.98
48	0.082	47.91	0.000	0.202	4.97
49	0.122	48.89	0.000	0.202	4.95
50	0.163	49.88	0.000	0.203	4.92
51	0.203	50.86	0.000	0.204	4.90
52	0.245	51.84	0.000	0.206	4.87
53	0.286	52.82	0.000	0.207	4.83
54	0.329	53.81	0.000	0.209	4.80
55	0.372	54.79	0.000	0.211	4.75
56	0.416	55.78	0.000	0.213	4.71
57	0.461	56.76	0.000	0.215	4.66
58	0.507	57.75	0.000	0.218	4.61
59	0.554	58.73	0.000	0.221	4.55
60	0.602	59.72	0.000	0.224	4.49
61	0.651	60.70	0.000	0.227	4.43
62	0.702	61.69	0.000	0.231	4.36
63	0.755	62.67	0.000	0.235	4.28
64	0.810	63.65	0.000	0.240	4.21
65	0.866	64.63	0.000	0.245	4.12
66	0.925	65.62	0.000	0.251	4.03
67	0.987	66.60	0.000	0.258	3.94
68	1.052	67.58	0.000	0.265	3.83
69	1.121	68.56	0.001	0.274	3.72
70	1.194	69.54	0.001	0.284	3.60
71	1.273	70.53	0.001	0.295	3.47
72	1.358	71.51	0.001	0.309	3.33
73	1.450	72.49	0.001	0.325	3.18
74	1.553	73.48	0.002	0.345	3.02
75	1.668	74.47	0.002	0.370	2.84
76	1.800	75.46	0.003	0.401	2.65
77	1.954	76.45	0.003	0.440	2.43
78	2.141	77.45	0.003	0.491	2.20
79	2.377	78.46	0.002	0.557	1.93
80	2.697	79.46	-0.009	0.634	1.62
81	3.191	80.48	-0.067	0.703	1.25
82	4.270	81.49	-0.481	0.778	0.72

Observed subscore:
A5+A6+A9+A10+A11+A12+A14+A17+A28+A29+A31+B5+B9+B11+B14+B15+C2+C6

	WML estimates			ML estimates				Expected	
Score	Theta	Bias	SEM	Theta	Bias	SEM		missing subscore	total score
0	-4.096	0.480	0.624	-4.773	0.328	0.627	0.0	0.000	0
1	-3.001	0.068	0.737	-3.381	-0.275	0.925	1.9	2.926	3
2	-2.453	0.010	0.679	-2.658	-0.262	0.890	3.7	5.717	6
3	-2.075	-0.001	0.601	-2.217	-0.192	0.769	5.5	8.485	8
4	-1.788	-0.003	0.537	-1.893	-0.140	0.659	7.3	11.254	11
5	-1.556	-0.003	0.489	-1.636	-0.104	0.576	9.0	14.038	14
6	-1.358	-0.002	0.454	-1.419	-0.078	0.517	10.8	16.849	17
7	-1.181	-0.001	0.428	-1.228	-0.060	0.474	12.7	19.700	20
8	-1.020	-0.001	0.408	-1.057	-0.045	0.444	14.6	22.602	23
9	-0.869	-0.001	0.394	-0.897	-0.034	0.422	16.6	25.564	26
10	-0.726	-0.000	0.383	-0.748	-0.025	0.407	18.6	28.593	29
11	-0.589	-0.000	0.376	-0.604	-0.018	0.396	20.7	31.693	32
12	-0.457	0.000	0.371	-0.466	-0.011	0.389	22.9	34.867	35
13	-0.326	0.000	0.369	-0.329	-0.004	0.385	25.1	38.112	38
14	-0.196	0.001	0.369	-0.194	0.002	0.384	27.4	41.421	41
15	-0.065	0.001	0.370	-0.058	0.007	0.385	29.8	44.777	45
16	0.068	0.000	0.373	0.080	0.013	0.389	32.2	48.160	48
17	0.205	-0.000	0.378	0.221	0.018	0.397	34.5	51.539	52
18	0.347	-0.001	0.385	0.368	0.025	0.410	36.9	54.884	55

19	0.494	-0.001	0.396	0.520	0.035	0.431	39.2	58.167	58
20	0.647	-0.001	0.412	0.681	0.049	0.463	41.4	61.364	61
21	0.808	0.001	0.435	0.854	0.071	0.510	43.5	64.468	64
22	0.979	0.003	0.466	1.045	0.104	0.574	45.5	67.484	67
23	1.171	0.006	0.508	1.266	0.151	0.658	47.4	70.434	70
24	1.400	0.007	0.562	1.542	0.214	0.761	49.3	73.348	73
25	1.703	-0.002	0.630	1.929	0.289	0.867	51.3	76.258	76
26	2.181	-0.060	0.695	2.602	0.295	0.898	53.2	79.168	79
27	3.278	-0.484	0.615	3.981	-0.324	0.617	55.0	82.000	82

Figur 7.1 Log-lineær Rasch model for MP4A

Observed subscore:
A2+A7+A8+A13+A15+A16+A18+A23+A24+A25+A32+B2+B6+B8+B10+B13+C3+C4

	WML estimates			ML estimates			Expected	missing	total
Score	Theta	Bias	SEM	Theta	Bias	SEM		subscore	score
0	-4.304	0.494	0.650	-4.957	0.339	0.644	0.0	0.000	0
1	-3.141	0.065	0.764	-3.525	-0.273	0.943	1.6	2.553	3
2	-2.570	0.009	0.699	-2.774	-0.260	0.902	3.1	5.143	5
3	-2.177	-0.001	0.617	-2.316	-0.190	0.777	4.8	7.772	8
4	-1.874	-0.002	0.549	-1.979	-0.138	0.664	6.5	10.451	10
5	-1.627	-0.002	0.499	-1.709	-0.104	0.579	8.2	13.186	13
6	-1.415	-0.001	0.460	-1.483	-0.082	0.518	10.0	15.969	16
7	-1.231	-0.001	0.431	-1.287	-0.066	0.474	11.8	18.789	19
8	-1.065	-0.001	0.408	-1.112	-0.054	0.442	13.6	21.633	22
9	-0.914	-0.002	0.389	-0.954	-0.045	0.417	15.5	24.489	24
10	-0.774	-0.002	0.374	-0.808	-0.037	0.399	17.3	27.349	27
11	-0.644	-0.002	0.363	-0.672	-0.030	0.386	19.2	30.208	30
12	-0.523	-0.002	0.355	-0.544	-0.022	0.377	21.1	33.064	33
13	-0.407	-0.002	0.349	-0.421	-0.014	0.372	22.9	35.918	36
14	-0.295	-0.001	0.346	-0.302	-0.005	0.371	24.8	38.779	39
15	-0.186	-0.001	0.346	-0.184	0.004	0.373	26.7	41.659	42
16	-0.077	0.000	0.349	-0.066	0.015	0.378	28.6	44.573	45
17	0.034	0.001	0.355	0.055	0.026	0.388	30.5	47.540	48
18	0.150	0.002	0.365	0.181	0.038	0.403	32.6	50.584	51

19	0.274	0.003	0.379	0.316	0.052	0.425	34.7	53.727	54
20	0.410	0.004	0.399	0.464	0.068	0.456	37.0	56.991	57
21	0.562	0.004	0.426	0.630	0.088	0.501	39.4	60.394	60
22	0.738	0.004	0.463	0.823	0.114	0.565	41.9	63.942	64
23	0.945	0.004	0.512	1.054	0.153	0.651	44.6	67.621	68
24	1.201	0.003	0.576	1.348	0.208	0.760	47.4	71.375	71
25	1.538	-0.008	0.651	1.756	0.278	0.875	50.1	75.104	75
26	2.047	-0.065	0.713	2.450	0.287	0.909	52.7	78.680	79
27	3.139	-0.482	0.617	3.835	-0.326	0.621	55.0	82.000	82

Figur 7.2 Log-lineær Rach model for MP4B

```
Observed subscore:
A1+A3+A4+A19+A20+A21+A22+A26+A27+A30+B1+B3+B4+B7+B12+B16+C1+C5


Expected
           WML estimates           ML estimates
missing   total
Score   Theta    Bias      SEM   Theta    Bias      SEM
subscore   score
--------------------------------------------------------------
-----------
  0    -3.984   0.517   0.694  -4.594   0.353   0.666    0.0
0.000   0
  1    -2.705   0.063   0.796  -3.109  -0.274   0.959    2.8
3.777   4
  2    -2.101   0.004   0.701  -2.332  -0.273   0.901    5.7
7.659   8
  3    -1.720  -0.005   0.607  -1.878  -0.208   0.775    8.4
11.406  11
  4    -1.443  -0.005   0.534  -1.558  -0.155   0.664   11.0
14.990  15
  5    -1.223  -0.004   0.481  -1.310  -0.117   0.578   13.4
18.429  18
  6    -1.036  -0.003   0.442  -1.105  -0.090   0.513   15.7
21.748  22
  7    -0.873  -0.002   0.412  -0.928  -0.070   0.466   18.0
24.968  25
  8    -0.727  -0.002   0.390  -0.771  -0.055   0.431   20.1
28.104  28
  9    -0.592  -0.001   0.372  -0.628  -0.044   0.405   22.2
31.169  31
 10    -0.466  -0.001   0.359  -0.495  -0.034   0.387   24.2
34.170  34
 11    -0.348  -0.001   0.349  -0.371  -0.027   0.373   26.1
37.111  37
 12    -0.235  -0.001   0.342  -0.252  -0.019   0.363   28.0
39.993  40
 13    -0.126  -0.001   0.337  -0.137  -0.013   0.357   29.8
42.816  43
 14    -0.020  -0.001   0.334  -0.026  -0.006   0.354   31.6
45.579  46
 15     0.085  -0.000   0.333   0.085   0.001   0.354   33.3
48.283  48
 16     0.190  -0.000   0.335   0.196   0.008   0.357   34.9
50.933  51
 17     0.295   0.000   0.338   0.308   0.016   0.363   36.5
53.535  54
 18     0.403   0.001   0.344   0.423   0.025   0.373   38.1
56.097  56
 19     0.515   0.001   0.353   0.542   0.035   0.389   39.6
58.631  59
```

20	0.634	0.001	0.366	0.669	0.047	0.413	41.1	61.147	61
21	0.760	0.002	0.384	0.806	0.063	0.446	42.7	63.656	64
22	0.898	0.003	0.408	0.958	0.085	0.493	44.2	66.170	66
23	1.053	0.004	0.440	1.131	0.115	0.558	45.7	68.696	69
24	1.232	0.005	0.483	1.335	0.158	0.643	47.2	71.240	71
25	1.451	0.003	0.536	1.594	0.218	0.746	48.8	73.807	74
26	1.735	-0.008	0.595	1.954	0.292	0.846	50.4	76.411	76
27	2.166	-0.066	0.640	2.578	0.300	0.863	52.1	79.096	79
28	3.123	-0.453	0.558	3.884	-0.304	0.586	54.0	82.000	82

Figur 7.3 Log-lineær Rasch model for MP4C

7.3 SeSam4

Tabel 4.2 Oversigt over forventede MP scores

Score på SeSam4 A-C	SeSam4A		SeSam4B		SeSam4C	
	E(MP4A)	E(MP4)	E(MP4B)	E(MP4)	E(MP4C)	E(MP4)
0	0	0	0	0	0	0
1	0.77	2.25	0.97	2.48	0.87	3.29
2	1.58	4.54	1.94	4.99	1.85	7.08
3	2.42	6.88	2.92	7.56	2.90	11.03
4	3.39	9.29	3.91	10.21	3.98	14.92
5	4.19	11.78	4.90	12.91	5.07	18.66
6	5.12	14.38	5.91	15.72	6.17	22.30
7	6.08	17.08	6.93	18.59	7.26	25.78
8	7.07	19.90	7.96	21.52	8.34	29.15
9	8.08	22.84	9.01	24.52	9.40	32.37
10	9.11	25.90	10.06	27.52	10.46	35.52
11	10.17	29.12	11.12	30.55	11.51	38.58
12	11.24	32.46	12.19	33.61	12.55	41.55
13	12.33	35.94	13.26	36.66	13.58	44.42
14	13.44	39.57	14.32	39.70	14.60	47.20
15	14.55	43.27	15.38	42.77	15.60	49.87
16	15.68	47.08	16.41	45.79	16.60	52.49
17	16.81	50.90	17.43	48.85	17.58	55.02
18	17.95	54.72	18.43	51.94	18.54	57.46
19	19.10	58.49	19.41	55.06	19.50	59.89
20	20.25	62.14	20.37	58.25	20.44	62.25
21	21.39	65.64	21.31	61.49	21.36	64.56
22	22.48	68.90	22.24	64.82	22.27	66.85
23	23.51	71.92	23.17	68.26	23.18	69.15
24	24.45	74.66	24.11	71.79	24.09	71.47
25	25.32	77.19	25.07	75.35	25.01	73.83
26	26.15	79.59	26.04	78.81	25.96	76.31
27	27	82	27	82	26.94	78.94
28					28	82

Endnu engang kun behov for begrænsede justeringer af SaSam4s deltest resultater i forhold til MP4s delskalaer.

7.4 De gentagne SeSam4 målinger

Analysen af de gentagne inkluderer kun elever, der har deltaget i alle tre test.

Tabel 7.3 Gennemsnitlige testresultater målt på den samlede MP skala

Skole	n	SeSam4A E(MP4)	SeSam4B E(MP4)	SESEAM4C E(MP4)	Parvis t-test A & B	Parvis t-test B & C
N	21	36.2	37.2	40.5	0.24	0.15
S	17	38.2	32.0	35.9	0.97	0.79
H	21	37.6	40.0	44.6	0.87	0.60
R	-	-	-	-	-	-
I alt	59	37.3	36.7	40.7	0.32	0.010

Der kan ikke påvises nogen ændringer fra SeSam4A til SeSam4B og kun en svagt signifikant ændring fra SeSam4B til SeSam4C, der ikke kan ses på de enkelte skoler

Tabel 7.4 Procentvis fordeling af elever ifølge Matematikprofilens kriteriebaserede kategorier

Skole	SeSam4	Begyndt	I gang	Godt i gang	Længere end forventet	Meget længere end forventet
N	SeSam4A	9.5	52.4	28.6	9.5	
	SeSam4B	4.8	66.7	19.0	9.5	
	SeSam4C	9.5	57.1	9.5	19.0	4.8
S	SeSam4A	17.6	35.3	35.3	5.9	5.9
	SeSam4B	11.8	76.5	11.8		
	SeSam4C	11.8	58.8	17.6	11.8	
H	SeSam4A	19.0	47.6	9.5	23.8	
	SeSam4B	14.3	42.9	28.6	14.3	
	SeSam4C	9.5	33.3	33.3	14.3	9.5
R	SeSam4A					
	SeSam4B					
	SeSam4C					
I alt	SeSam4A	15.3	45.8	23.7	13.6	1.7
	SeSam4B	10.2	61.0	20.3	8.5	
	SeSam4C	10.2	49.2	20.3	15.3	5.1

8.SeSam5

8.1 Matematikprofilen 5

Tabel 8.1 WML estimater af personparameteren i MP5.

Score	Theta estimate	True score	Bias	SEM	Score SEM
0	-5.581	0.47	0.500	0.830	0.67
1	-4.389	1.43	0.066	0.778	1.14
2	-3.803	2.41	0.009	0.703	1.46
3	-3.405	3.41	0.000	0.616	1.71
4	-3.104	4.41	-0.001	0.544	1.93
5	-2.862	5.41	-0.001	0.490	2.14
6	-2.660	6.41	-0.001	0.448	2.32
7	-2.487	7.41	-0.001	0.415	2.49
8	-2.336	8.40	-0.001	0.389	2.65
9	-2.201	9.40	-0.001	0.367	2.79
10	-2.079	10.40	-0.001	0.348	2.93
11	-1.969	11.39	-0.001	0.333	3.06
12	-1.867	12.39	-0.000	0.319	3.19
13	-1.773	13.38	-0.000	0.307	3.30
14	-1.685	14.37	-0.000	0.297	3.41
15	-1.602	15.36	-0.000	0.288	3.52
16	-1.524	16.35	-0.000	0.279	3.62
17	-1.450	17.35	-0.000	0.272	3.71
18	-1.380	18.34	-0.000	0.265	3.80
19	-1.313	19.33	-0.000	0.259	3.89
20	-1.249	20.32	-0.000	0.254	3.97
21	-1.187	21.31	-0.000	0.249	4.05
22	-1.128	22.30	-0.000	0.244	4.12
23	-1.071	23.29	-0.000	0.240	4.19
24	-1.015	24.28	-0.000	0.236	4.26
25	-0.961	25.27	-0.000	0.232	4.32
26	-0.909	26.26	-0.000	0.229	4.38
27	-0.858	27.25	-0.000	0.226	4.44
28	-0.809	28.24	-0.000	0.223	4.49
29	-0.760	29.23	-0.000	0.221	4.54
30	-0.713	30.21	-0.000	0.218	4.59
31	-0.667	31.20	-0.000	0.216	4.64
32	-0.621	32.19	-0.000	0.214	4.68
33	-0.576	33.17	-0.000	0.212	4.72
34	-0.532	34.16	-0.000	0.211	4.75
35	-0.489	35.14	-0.000	0.209	4.78
36	-0.446	36.13	-0.000	0.208	4.81
37	-0.404	37.11	-0.000	0.207	4.83
38	-0.362	38.09	-0.000	0.206	4.85
39	-0.321	39.07	-0.000	0.205	4.87
40	-0.279	40.06	-0.000	0.205	4.88

Tabellen fortsættes på næste side

Score	Theta estimate	True score	Bias	SEM	Score SEM
41	-0.238	41.04	-0.000	0.205	4.89
42	-0.197	42.02	-0.000	0.204	4.90
43	-0.157	43.00	-0.000	0.204	4.90
44	-0.116	43.98	-0.000	0.204	4.90
45	-0.075	44.96	0.000	0.205	4.89
46	-0.034	45.94	0.000	0.205	4.88
47	0.007	46.92	0.000	0.206	4.87
48	0.049	47.90	0.000	0.207	4.85
49	0.091	48.88	0.000	0.208	4.82
50	0.133	49.86	0.000	0.209	4.80
51	0.176	50.84	0.000	0.210	4.77
52	0.220	51.83	0.000	0.212	4.73
53	0.264	52.81	0.000	0.213	4.69
54	0.309	53.79	0.000	0.215	4.65
55	0.355	54.78	0.000	0.218	4.61
56	0.402	55.76	0.000	0.220	4.56
57	0.450	56.75	0.000	0.223	4.50
58	0.499	57.74	0.000	0.226	4.45
59	0.550	58.73	0.000	0.229	4.39
60	0.602	59.72	0.000	0.232	4.33
61	0.656	60.71	0.000	0.236	4.26
62	0.711	61.70	0.000	0.240	4.19
63	0.769	62.68	0.000	0.244	4.12
64	0.828	63.67	0.000	0.249	4.04
65	0.890	64.66	-0.000	0.255	3.96
66	0.955	65.65	-0.000	0.261	3.87
67	1.022	66.64	0.000	0.268	3.78
68	1.093	67.62	0.000	0.275	3.68
69	1.168	68.61	0.000	0.284	3.58
70	1.247	69.59	0.000	0.294	3.47
71	1.332	70.57	0.001	0.306	3.34
72	1.423	71.56	0.001	0.320	3.21
73	1.522	72.54	0.001	0.336	3.07
74	1.632	73.52	0.002	0.356	2.92
75	1.754	74.50	0.002	0.381	2.75
76	1.894	75.49	0.003	0.413	2.56
77	2.059	76.48	0.003	0.454	2.35
78	2.259	77.48	0.004	0.508	2.12
79	2.515	78.49	0.002	0.579	1.86
80	2.865	79.50	-0.008	0.666	1.55
81	3.412	80.52	-0.066	0.748	1.19
82	4.579	81.52	-0.498	0.818	0.68

Observed subscore:
A1+A2+A3+A4+A5+A6+A15+A17+A18+A28+A29+B1+B7+B12+B13+B16+C1+C4

	WML estimates			ML estimates			Expected		
Score	Theta	Bias	SEM	Theta	Bias	SEM	missing subscore		total score
0	-4.175	0.518	0.694	-4.786	0.354	0.668	0.0	0.000	0
1	-2.897	0.064	0.801	-3.297	-0.272	0.963	2.7	3.738	4
2	-2.286	0.006	0.708	-2.513	-0.270	0.905	5.2	7.249	7
3	-1.893	-0.003	0.612	-2.050	-0.206	0.777	7.6	10.649	11
4	-1.606	-0.004	0.538	-1.723	-0.154	0.664	9.9	13.924	14
5	-1.378	-0.004	0.483	-1.470	-0.118	0.577	12.1	17.077	17
6	-1.188	-0.003	0.442	-1.262	-0.093	0.512	14.1	20.116	20
7	-1.023	-0.003	0.411	-1.084	-0.075	0.465	16.1	23.054	23
8	-0.877	-0.002	0.386	-0.928	-0.060	0.429	17.9	25.904	26
9	-0.745	-0.002	0.367	-0.787	-0.049	0.403	19.7	28.674	29
10	-0.624	-0.002	0.353	-0.658	-0.039	0.384	21.4	31.376	31
11	-0.512	-0.002	0.342	-0.539	-0.030	0.370	23.0	34.017	34
12	-0.406	-0.002	0.334	-0.426	-0.022	0.360	24.6	36.607	37
13	-0.305	-0.002	0.329	-0.317	-0.013	0.355	26.2	39.158	39
14	-0.207	-0.001	0.327	-0.211	-0.004	0.353	27.7	41.680	42
15	-0.110	-0.000	0.327	-0.107	0.005	0.356	29.2	44.189	44
16	-0.013	0.000	0.330	-0.002	0.016	0.362	30.7	46.700	47
17	0.086	0.001	0.336	0.106	0.027	0.373	32.2	49.230	49
18	0.188	0.002	0.346	0.219	0.040	0.389	33.8	51.802	52
19	0.298	0.003	0.360	0.339	0.055	0.413	35.4	54.441	54
20	0.417	0.004	0.379	0.471	0.072	0.446	37.2	57.178	57
21	0.550	0.005	0.405	0.620	0.094	0.492	39.1	60.050	60
22	0.704	0.005	0.440	0.793	0.122	0.556	41.1	63.101	63
23	0.889	0.005	0.487	1.004	0.161	0.641	43.4	66.377	66
24	1.122	0.002	0.548	1.274	0.217	0.748	45.9	69.916	70
25	1.434	-0.010	0.619	1.655	0.286	0.856	48.7	73.715	74
26	1.902	-0.069	0.672	2.308	0.294	0.882	51.7	77.693	78
27	2.911	-0.465	0.583	3.644	-0.312	0.599	55.0	82.000	82

Figur 8.1 Log-lineær Rasch model for MP5A

Observed subscore:
A9+A10+A11+A13+A19+A20+A22+A23+A25+A30+A32+B3+B4+B8+B14+B15+C2+C6

	WML estimates			ML estimates			Expected		
Score	Theta	Bias	SEM	Theta	Bias	SEM	missing subscore		total score
0	-4.200	0.492	0.643	-4.861	0.336	0.639	0.0	0.000	0
1	-3.048	0.063	0.749	-3.440	-0.278	0.933	2.3	3.306	3
2	-2.497	0.006	0.683	-2.708	-0.267	0.894	4.2	6.159	6
3	-2.128	-0.003	0.604	-2.267	-0.195	0.774	5.9	8.897	9
4	-1.846	-0.003	0.541	-1.945	-0.139	0.664	7.6	11.612	12
5	-1.612	-0.002	0.493	-1.687	-0.101	0.580	9.3	14.346	14
6	-1.408	-0.001	0.458	-1.467	-0.075	0.519	11.1	17.124	17
7	-1.226	-0.001	0.432	-1.272	-0.057	0.476	13.0	19.959	20
8	-1.058	-0.001	0.412	-1.095	-0.044	0.446	14.9	22.856	23
9	-0.902	-0.001	0.398	-0.932	-0.034	0.425	16.8	25.819	26
10	-0.756	-0.001	0.387	-0.779	-0.025	0.410	18.8	28.847	29
11	-0.617	-0.001	0.380	-0.632	-0.017	0.401	20.9	31.941	32
12	-0.482	-0.000	0.376	-0.491	-0.010	0.395	23.1	35.102	35
13	-0.351	0.000	0.375	-0.352	-0.002	0.393	25.3	38.334	38
14	-0.219	0.001	0.375	-0.213	0.005	0.392	27.6	41.636	42
15	-0.086	0.002	0.378	-0.073	0.012	0.394	30.0	45.003	45
16	0.053	0.002	0.383	0.071	0.018	0.398	32.4	48.416	48
17	0.198	0.001	0.389	0.221	0.023	0.405	34.8	51.848	52
18	0.352	0.000	0.398	0.378	0.027	0.416	37.3	55.255	55
19	0.515	-0.001	0.409	0.543	0.033	0.435	39.6	58.590	59
20	0.686	-0.002	0.424	0.718	0.043	0.464	41.8	61.810	62
21	0.866	-0.002	0.445	0.904	0.060	0.508	43.9	64.887	65
22	1.054	-0.001	0.473	1.107	0.089	0.572	45.8	67.816	68
23	1.257	0.002	0.509	1.335	0.135	0.658	47.6	70.614	71
24	1.487	0.003	0.554	1.609	0.203	0.760	49.3	73.322	73
25	1.772	-0.006	0.604	1.977	0.286	0.856	51.0	76.006	76
26	2.194	-0.062	0.639	2.603	0.300	0.868	52.8	78.778	79
27	3.153	-0.454	0.557	3.918	-0.306	0.589	55.0	82.000	82

Figur 8.2 Log-lineær Rach model for MP5B

Observed subscore:
A7+A8+A12+A14+A16+A21+A24+A26+A27+A31+B2+B5+B6+B9+B10+B11+C3+C5

	WML estimates			ML estimates			Expected		
							missing		total
Score	Theta	Bias	SEM	Theta	Bias	SEM	subscore		score
0	-4.954	0.542	0.758	-5.484	0.383	0.714	0.0	0.000	0
1	-3.519	0.065	0.891	-3.892	-0.248	1.022	1.2	2.227	2
2	-2.775	0.007	0.781	-2.993	-0.249	0.946	2.8	4.840	5
3	-2.290	-0.003	0.670	-2.445	-0.190	0.806	4.7	7.675	8
4	-1.935	-0.004	0.589	-2.051	-0.143	0.691	6.6	10.639	11
5	-1.654	-0.003	0.530	-1.745	-0.110	0.606	8.7	13.685	14
6	-1.420	-0.003	0.487	-1.492	-0.086	0.544	10.8	16.784	17
7	-1.217	-0.002	0.454	-1.274	-0.068	0.499	12.9	19.923	20
8	-1.036	-0.002	0.430	-1.082	-0.054	0.466	15.1	23.096	23
9	-0.871	-0.001	0.412	-0.907	-0.042	0.441	17.3	26.305	26
10	-0.716	-0.001	0.398	-0.745	-0.033	0.422	19.6	29.552	30
11	-0.570	-0.000	0.387	-0.592	-0.025	0.407	21.8	32.835	33
12	-0.430	-0.000	0.379	-0.445	-0.018	0.397	24.1	36.146	36
13	-0.293	-0.000	0.374	-0.304	-0.012	0.390	26.5	39.469	39
14	-0.160	-0.000	0.370	-0.166	-0.006	0.386	28.8	42.779	43
15	-0.028	-0.001	0.369	-0.029	-0.001	0.385	31.0	46.049	46
16	0.104	-0.001	0.370	0.107	0.006	0.388	33.3	49.253	49
17	0.235	-0.001	0.373	0.244	0.013	0.394	35.4	52.371	52
18	0.368	-0.000	0.379	0.384	0.021	0.404	37.4	55.395	55
19	0.505	0.000	0.388	0.529	0.032	0.420	39.3	58.326	58
20	0.648	0.001	0.401	0.682	0.045	0.443	41.2	61.177	61
21	0.800	0.002	0.421	0.846	0.062	0.476	43.0	63.968	64
22	0.966	0.003	0.448	1.028	0.084	0.522	44.7	66.721	67
23	1.153	0.005	0.486	1.236	0.114	0.586	46.5	69.455	69
24	1.372	0.007	0.539	1.485	0.157	0.672	48.2	72.177	72
25	1.646	0.008	0.612	1.805	0.213	0.783	49.9	74.874	75
26	2.023	-0.000	0.711	2.263	0.280	0.909	51.5	77.498	77
27	2.644	-0.058	0.823	3.064	0.276	0.975	52.9	79.938	80
28	4.028	-0.538	0.731	4.601	-0.367	0.687	54.0	82.000	82

Figur 8.3 Log-lineær Rasch model for MP5C

8.3 SeSam5

Tabel 8.2 Oversigt over forventede MP scores

Score på SeSam5 A-C	SeSam5A		SeSam5B		SeSam5C	
	E(MP5A)	E(MP5)	E(MP5B)	E(MP5)	E(MP5C)	E(MP5)
0	0	0	0	0	0	0
1	1.00	3.74	1.05	3.44	0.81	1.80
2	2.02	7.32	2.07	6.35	1.72	4.11
3	3.06	10.85	3.07	9.09	2.71	6.85
4	4.11	14.27	4.05	11.75	3.76	9.93
5	5.18	17.62	5.03	14.43	4.82	13.14
6	6.27	20.91	6.01	17.15	5.88	16.41
7	7.37	24.11	6.99	19.93	6.94	19.74
8	8.48	27.23	7.97	22.77	8.00	23.10
9	9.60	30.29	8.97	25.73	9.07	26.53
10	10.72	33.28	9.98	28.79	10.13	29.98
11	11.84	36.19	11.00	31.94	11.19	33.46
12	12.97	39.08	12.03	35.20	12.24	36.94
13	14.08	41.88	13.06	38.53	13.30	40.46
14	15.17	44.62	14.10	41.97	14.35	43.92
15	16.24	47.31	15.15	45.52	15.39	47.30
16	17.29	49.98	16.20	49.10	16.42	50.56
17	18.30	52.59	17.27	52.77	17.45	53.73
18	19.28	55.21	18.36	56.46	18.47	56.77
19	20.22	57.81	19.47	60.10	19.48	59.69
20	21.14	60.48	20.60	63.66	20.48	62.52
21	22.04	63.23	21.73	67.02	21.46	65.23
22	22.93	66.15	22.81	70.08	22.44	67.92
23	23.82	69.28	23.81	72.81	23.40	70.54
24	24.70	72.58	24.68	75.15	24.35	73.12
25	25.55	75.90	24.45	77.25	25.29	75.64
26	26.33	79.11	26.18	79.36	26.20	77.99
27	27	82	27	82	27.10	80.14
28					28	82

God overensstemmelse mellem SeSam5 deltest og MP5s delskalaer.

8.4 De gentagne SeSam5 målinger

Analysen af de gentagne inkluderer kun elever, der har deltaget i alle tre test.

Tabel 8.3 Gennemsnitlige testresultater målt på den samlede MP skala

Skole	n	SeSam5A E(MP5)	SeSam5B E(MP5)	SESEAM5C E(MP5)	Parvis t-test A & B	Parvis t-test B & C
N	20	46.5	45.4	48.5	0.27	0.17
S	22	34.8	27.5	24.4	0.98	0.12
H	22	55.1	41.2	43.7	*	0.20
R	44	39.3	38.4	39.6	0.27	0.20
I alt						

Ingen påviselige ændringer fra SeSam5A til SeSam5 B og C. SeSam5A resultaterne på skolerne N og S er påfaldende høje. Det giver ingen mening at beregne et samlet resultat.

Tabel 8.4 Procentvis fordeling af elever ifølge Matematikprofilens kriteriebaserede kategorier

Skole	SeSam5	Begyndt	I gang	Godt i gang	Længere end forventet	Meget længere end forventet
N	SeSam5A	5.0	40.0	25.0	25.0	5.0
	SeSam5B	5.0	40.0	15.0	35.0	5.0
	SeSam5C	10.0	25.0	25.0	20.0	20.0
S	SeSam5A	31.8	36.4	18.2	9.1	4.5
	SeSam5B	40.9	45.5	9.1	4.5	
	SeSam5C	40.9	40.9	18.2		
H	SeSam5A	*	*	*	*	*
	SeSam5B	13.6	40.9	22.7	22.7	
	SeSam5C		50.0	36.4	9.1	4.5
R	SeSam5A	6.8	45.5	40.9	4.5	2.3
	SeSam5B	4.5	63.6	22.7	6.8	2.3
	SeSam5C	9.1	45.5	31.8	11.4	2.3
I alt	SeSam5A					
	SeSam5B	13.9	50.9	18.5	14.8	1.9
	SeSam5C	13.9	41.7	28.7	10.2	5.6

9. SeSam6

9.1 Matematikprofilen 6

Tabel 6.1 WML estimater af personparameteren i MP6.

Score	Theta estimate	True score	Bias	SEM	Score SEM
0	-5.425	0.48	0.491	0.809	0.69
1	-4.279	1.44	0.066	0.755	1.17
2	-3.721	2.42	0.009	0.688	1.48
3	-3.341	3.40	0.000	0.607	1.73
4	-3.048	4.39	-0.001	0.539	1.95
5	-2.810	5.39	-0.001	0.487	2.14
6	-2.609	6.39	-0.001	0.447	2.31
7	-2.435	7.38	-0.001	0.416	2.47
8	-2.281	8.38	-0.001	0.391	2.62
9	-2.143	9.38	-0.001	0.370	2.76
10	-2.019	10.37	-0.001	0.352	2.90
11	-1.905	11.37	-0.001	0.337	3.02
12	-1.800	12.37	-0.000	0.323	3.14
13	-1.703	13.36	-0.000	0.312	3.25
14	-1.612	14.35	-0.000	0.301	3.36
15	-1.527	15.34	-0.000	0.292	3.46
16	-1.447	16.33	-0.000	0.284	3.56
17	-1.370	17.33	-0.000	0.277	3.65
18	-1.297	18.32	-0.000	0.270	3.73
19	-1.228	19.31	-0.000	0.264	3.82
20	-1.161	20.30	-0.000	0.258	3.89
21	-1.097	21.29	-0.000	0.253	3.97
22	-1.035	22.28	-0.000	0.249	4.04
23	-0.976	23.27	-0.000	0.245	4.10
24	-0.918	24.26	-0.000	0.241	4.17
25	-0.862	25.25	-0.000	0.237	4.23
26	-0.807	26.24	-0.000	0.234	4.28
27	-0.754	27.23	-0.000	0.231	4.34
28	-0.702	28.22	-0.000	0.229	4.39
29	-0.651	29.20	-0.000	0.226	4.43
30	-0.601	30.19	-0.000	0.224	4.48
31	-0.552	31.18	-0.000	0.222	4.52
32	-0.505	32.17	-0.000	0.220	4.56
33	-0.457	33.15	-0.000	0.218	4.59
34	-0.411	34.14	-0.000	0.217	4.62
35	-0.365	35.12	-0.000	0.215	4.65
36	-0.320	36.11	-0.000	0.214	4.67
37	-0.275	37.10	-0.000	0.213	4.70
38	-0.230	38.08	-0.000	0.212	4.71
39	-0.186	39.07	-0.000	0.212	4.73
40	-0.142	40.05	-0.000	0.211	4.74

Tabellen fortsættes på næste side

Score	Theta estimate	True score	Bias	SEM	Score SEM
41	-0.099	41.03	-0.000	0.211	4.75
42	-0.055	42.02	-0.000	0.210	4.75
43	-0.012	43.00	0.000	0.210	4.76
44	0.032	43.98	0.000	0.210	4.75
45	0.076	44.97	0.000	0.211	4.75
46	0.119	45.95	0.000	0.211	4.74
47	0.163	46.93	0.000	0.212	4.73
48	0.207	47.92	0.000	0.212	4.71
49	0.252	48.90	0.000	0.213	4.70
50	0.297	49.89	0.000	0.214	4.67
51	0.342	50.87	0.000	0.215	4.65
52	0.388	51.86	0.000	0.217	4.62
53	0.434	52.85	0.000	0.218	4.59
54	0.482	53.84	0.000	0.220	4.56
55	0.530	54.82	0.000	0.222	4.52
56	0.578	55.81	0.000	0.224	4.48
57	0.628	56.80	0.000	0.226	4.43
58	0.679	57.79	0.000	0.228	4.39
59	0.731	58.78	0.000	0.231	4.34
60	0.784	59.77	-0.000	0.234	4.29
61	0.839	60.76	-0.000	0.237	4.23
62	0.895	61.75	-0.000	0.240	4.18
63	0.952	62.73	-0.000	0.244	4.11
64	1.011	63.72	-0.000	0.248	4.05
65	1.073	64.71	-0.000	0.253	3.98
66	1.136	65.69	-0.000	0.258	3.90
67	1.202	66.68	-0.000	0.264	3.82
68	1.270	67.66	-0.000	0.271	3.74
69	1.342	68.64	0.000	0.278	3.64
70	1.418	69.61	0.000	0.287	3.54
71	1.498	70.59	0.000	0.298	3.43
72	1.584	71.56	0.001	0.311	3.31
73	1.677	72.54	0.001	0.326	3.17
74	1.779	73.51	0.002	0.345	3.02
75	1.892	74.49	0.002	0.368	2.86
76	2.020	75.47	0.003	0.398	2.67
77	2.169	76.45	0.004	0.436	2.46
78	2.349	77.44	0.004	0.485	2.23
79	2.576	78.44	0.003	0.549	1.96
80	2.886	79.44	-0.008	0.625	1.65
81	3.368	80.46	-0.066	0.694	1.27
82	4.437	81.49	-0.479	0.773	0.72

Observed subscore:
A2+A9+A10+A12+A18+A19+A21+A23+A27+A30+A32+B4+B6+B11+B13+B14+C4+C6

	WML estimates			ML estimates			Expected		
Score	Theta	Bias	SEM	Theta	Bias	SEM	missing subscore		total score
0	-4.610	0.506	0.683	-5.218	0.352	0.666	0.0	0.000	0
1	-3.369	0.064	0.810	-3.739	-0.261	0.973	1.3	2.327	2
2	-2.735	0.008	0.735	-2.935	-0.248	0.922	2.7	4.735	5
3	-2.297	-0.002	0.646	-2.434	-0.180	0.791	4.2	7.217	7
4	-1.962	-0.003	0.576	-2.064	-0.130	0.680	5.8	9.770	10
5	-1.690	-0.002	0.526	-1.767	-0.096	0.599	7.4	12.408	12
6	-1.456	-0.001	0.490	-1.514	-0.072	0.542	9.1	15.149	15
7	-1.246	-0.000	0.463	-1.291	-0.055	0.501	11.0	18.009	18
8	-1.053	-0.000	0.442	-1.089	-0.043	0.471	13.0	20.999	21
9	-0.871	-0.000	0.427	-0.901	-0.034	0.449	15.1	24.121	24
10	-0.700	-0.000	0.415	-0.724	-0.027	0.433	17.4	27.365	27
11	-0.536	-0.000	0.406	-0.556	-0.021	0.421	19.7	30.711	31
12	-0.380	-0.001	0.399	-0.394	-0.015	0.413	22.1	34.133	34
13	-0.227	-0.001	0.395	-0.237	-0.010	0.409	24.6	37.599	38
14	-0.078	-0.000	0.392	-0.083	-0.004	0.406	27.1	41.080	41
15	0.069	-0.000	0.392	0.070	0.002	0.406	29.5	44.545	45
16	0.217	0.000	0.393	0.223	0.008	0.409	32.0	47.972	48
17	0.366	0.000	0.396	0.379	0.014	0.414	34.3	51.345	51
18	0.520	-0.000	0.402	0.538	0.021	0.424	36.7	54.653	55
19	0.680	-0.001	0.411	0.704	0.030	0.441	38.9	57.893	58
20	0.847	-0.001	0.425	0.878	0.043	0.469	41.1	61.065	61
21	1.023	-0.001	0.445	1.064	0.062	0.511	43.2	64.171	64
22	1.210	0.001	0.474	1.267	0.091	0.574	45.2	67.217	67
23	1.416	0.003	0.513	1.498	0.138	0.660	47.2	70.212	70
24	1.651	0.004	0.563	1.777	0.204	0.763	49.2	73.173	73
25	1.946	-0.004	0.618	2.154	0.285	0.863	51.1	76.116	76
26	2.387	-0.059	0.663	2.800	0.298	0.882	53.1	79.060	79
27	3.404	-0.466	0.580	4.144	-0.314	0.601	55.0	82.000	82

Figur 9.1 Log-lineær Rasch model for MP6A

Observed subscore:
A4+A5+A11+A13+A14+A16+A20+A24+A25+A29+B1+B2+B3+B8+B9+B12+C1+C2

	WML estimates			ML estimates			Expected		
Score	Theta	Bias	SEM	Theta	Bias	SEM	missing subscore		total score
0	-3.229	0.462	0.573	-3.978	0.310	0.595	0.0	0.000	0
1	-2.237	0.066	0.655	-2.651	-0.299	0.873	5.0	6.026	6
2	-1.793	0.008	0.606	-2.013	-0.291	0.852	8.2	10.181	10
3	-1.500	-0.003	0.543	-1.645	-0.218	0.749	10.7	13.667	14
4	-1.277	-0.004	0.487	-1.382	-0.158	0.645	12.8	16.797	17
5	-1.094	-0.004	0.444	-1.173	-0.116	0.559	14.7	19.703	20
6	-0.936	-0.003	0.410	-0.998	-0.087	0.494	16.5	22.459	22
7	-0.796	-0.002	0.385	-0.845	-0.066	0.446	18.1	25.108	25
8	-0.669	-0.002	0.366	-0.708	-0.050	0.412	19.7	27.679	28
9	-0.551	-0.001	0.352	-0.581	-0.038	0.388	21.2	30.193	30
10	-0.439	-0.001	0.342	-0.462	-0.028	0.371	22.7	32.666	33
11	-0.333	-0.001	0.335	-0.349	-0.020	0.360	24.1	35.109	35
12	-0.230	-0.000	0.330	-0.240	-0.012	0.353	25.5	37.532	38
13	-0.129	-0.000	0.328	-0.133	-0.004	0.349	26.9	39.945	40
14	-0.029	0.000	0.328	-0.026	0.003	0.348	28.4	42.358	42
15	0.072	0.001	0.330	0.081	0.010	0.351	29.8	44.779	45
16	0.175	0.001	0.334	0.189	0.017	0.356	31.2	47.217	47
17	0.281	0.001	0.340	0.302	0.024	0.364	32.7	49.685	50
18	0.391	0.001	0.348	0.419	0.032	0.376	34.2	52.193	52
19	0.508	0.002	0.360	0.543	0.041	0.393	35.8	54.755	55
20	0.634	0.002	0.375	0.677	0.052	0.416	37.4	57.387	57
21	0.772	0.002	0.395	0.824	0.066	0.449	39.1	60.106	60
22	0.924	0.002	0.422	0.988	0.084	0.496	40.9	62.931	63
23	1.096	0.002	0.456	1.175	0.111	0.561	42.9	65.882	66
24	1.295	0.003	0.502	1.398	0.151	0.647	45.0	68.969	69
25	1.536	0.002	0.559	1.676	0.209	0.754	47.2	72.186	72
26	1.847	-0.008	0.623	2.060	0.283	0.861	49.5	75.484	75
27	2.309	-0.065	0.671	2.713	0.294	0.883	51.8	78.775	79
28	3.317	-0.464	0.581	4.051	-0.313	0.600	54.0	82.000	82

Figur 9.2 Log-lineær Rach model for MP6B

Observed subscore:
A1+A3+A6+A7+A8+A15+A17+A22+A26+A28+A31+B5+B7+B10+B15+B16+C3+C5

Score	WML estimates Theta	WML estimates Bias	WML estimates SEM	ML estimates Theta	ML estimates Bias	ML estimates SEM	Expected missing subscore	Expected total score	
0	-4.620	0.504	0.681	-5.228	0.352	0.668	0.0	0.000	0
1	-3.388	0.065	0.816	-3.748	-0.257	0.979	1.3	2.309	2
2	-2.745	0.010	0.750	-2.933	-0.238	0.931	2.7	4.743	5
3	-2.286	0.002	0.663	-2.414	-0.169	0.798	4.3	7.340	7
4	-1.922	0.000	0.592	-2.021	-0.124	0.684	6.1	10.115	10
5	-1.620	-0.001	0.539	-1.703	-0.096	0.601	8.1	13.054	13
6	-1.365	-0.002	0.498	-1.434	-0.078	0.544	10.1	16.124	16
7	-1.143	-0.002	0.467	-1.202	-0.065	0.503	12.3	19.284	19
8	-0.947	-0.002	0.443	-0.996	-0.053	0.473	14.5	22.503	23
9	-0.770	-0.002	0.424	-0.809	-0.043	0.451	16.8	25.762	26
10	-0.608	-0.002	0.409	-0.638	-0.034	0.434	19.0	29.049	29
11	-0.455	-0.001	0.399	-0.477	-0.024	0.421	21.4	32.359	32
12	-0.310	-0.000	0.392	-0.323	-0.016	0.412	23.7	35.689	36
13	-0.167	0.000	0.388	-0.173	-0.008	0.406	26.0	39.035	39
14	-0.026	0.001	0.386	-0.025	-0.000	0.402	28.4	42.392	42
15	0.116	0.001	0.387	0.124	0.006	0.402	30.7	45.748	46
16	0.262	0.001	0.389	0.274	0.011	0.404	33.1	49.089	49
17	0.413	0.000	0.394	0.429	0.017	0.410	35.4	52.395	52
18	0.569	-0.000	0.401	0.588	0.022	0.421	37.6	55.646	56
19	0.731	-0.001	0.411	0.754	0.030	0.439	39.8	58.825	59
20	0.898	-0.002	0.424	0.928	0.041	0.467	41.9	61.920	62
21	1.074	-0.001	0.444	1.113	0.060	0.511	43.9	64.930	65
22	1.259	0.000	0.472	1.314	0.091	0.574	45.9	67.858	68
23	1.460	0.003	0.510	1.541	0.138	0.660	47.7	70.719	71
24	1.691	0.005	0.558	1.816	0.206	0.762	49.5	73.531	74
25	1.980	-0.004	0.612	2.188	0.287	0.860	51.3	76.326	76
26	2.411	-0.059	0.655	2.826	0.300	0.876	53.1	79.139	79
27	3.414	-0.463	0.574	4.160	-0.312	0.597	55.0	82.000	82

Figur 9.3 Log-lineær Rasch model for MP6C

9.3 SeSam6

Tabel 9.2 Oversigt over forventede MP scores

Score på SeSam6 A-C	SeSam6A		SeSam6B		SeSam6C	
	E(MP6A)	E(MP6)	E(MP6B)	E(MP6)	E(MP6C)	E(MP6)
0	0	0	0	0	0	0
1	0.68	1.61	1.03	5.87	0.92	2.13
2	1.38	3.30	2.08	10.06	1.85	4.40
3	2.12	5.13	3.12	13.58	2.80	6.87
4	2.95	7.24	4.16	16.78	3.76	9.52
5	3.87	9.64	5.20	19.79	4.75	12.42
6	4.85	12.27	6.24	22.66	5.75	15.48
7	5.89	15.16	7.28	25.44	6.76	18.65
8	6.94	18.20	8.31	28.13	7.79	21.95
9	8.00	21.39	9.34	30.78	8.83	25.31
10	9.06	24.71	10.37	33.39	9.87	28.68
11	10.11	28.10	11.40	35.97	10.92	32.10
12	11.17	31.60	12.42	38.50	11.97	35.54
13	12.23	35.16	13.44	41.02	13.03	39.05
14	13.30	38.79	14.44	43.48	14.08	42.55
15	14.40	42.52	15.45	45.97	15.14	46.10
16	15.52	46.30	16.44	48.43	16.19	49.62
17	16.68	50.17	17.42	50.88	17.25	53.15
18	17.88	54.09	18.40	53.38	18.30	56.58
19	19.13	58.09	19.37	55.90	19.35	59.93
20	20.45	62.20	20.33	58.46	20.39	63.16
21	21.85	66.48	21.28	61.08	21.42	66.25
22	23.30	70.82	22.24	63.83	22.44	6923
23	24.72	75.05	23.19	66.66	23.45	72.10
24	25.98	78.85	24.15	69.63	24.44	74.87
25	26.30	79.93	25.10	72.68	25.39	77.51
26	26.67	80.98	26.06	75.82	26.27	79.96
27	27	82	27.03	78.95	27	82
28			28	82		

God tilpasning mellem SeSam6s deltest og MP66s delskalaer.

9.4 De gentagne SeSam6 målinger

Analysen af de gentagne inkluderer kun elever, der har deltaget i alle tre test.

Tabel 9.3 Gennemsnitlige testresultater målt på den samlede MP skala

Skole	n	SeSam6A E(MP6)	SeSam6B E(MP6)	SESEAM6C E(MP6)	Parvis t-test A & B	Parvis t-test B & C
N	25	50.5	47.4	53.7	0.92	<0.0005
S	32	29.2	23.5	26.9	*	0.065
H						
R						
I alt						

Endnu engang problemer med for høje delskala resultater på SeSam6A. Især på skole S.

Tabel 9.4 Procentvis fordeling af elever ifølge Matematikprofilens kriteriebaserede kategorier

Skole	SeSam6	Begyndt	I gang	Godt i gang	Længere end forventet	Meget længere end forventet
N	SeSam6A		24.0	36.0	28.0	12.0
	SeSam6B		32.0	56.0	4.0	8.0
	SeSam6C		16.0	24.0	40.0	20.0
S	SeSam6A					
	SeSam6B	37.5	56.3	6.3		
	SeSam6C	31.3	59.4	9.4		
H	SeSam6A					
	SeSam6B					
	SeSam6C					
R	SeSam6A					
	SeSam6B					
	SeSam6C					
I alt	SeSam6A					
	SeSam6B					
	SeSam6C					

9.10 Appendix J: Conversion of SeSam score to MP score

SESAM-1A	MP1	SESAM-1B	MP1	SESAM1C	MP1

0	0	0	0	0	0
1	2	1	5	1	3
2	4	2	7	2	6
3	6	3	9	3	10
4	9	4	12	4	14
5	12	5	14	5	18
6	15	6	16	6	23
7	18	7	18	7	27
8	22	8	20	8	32
9	25	9	23	9	36
10	29	10	25	10	41
11	32	11	28	11	45
12	36	12	31	12	49
13	39	13	34	13	53
14	42	14	38	14	55
15	46	15	42	15	58
16	49	16	46	16	59
17	53	17	50	17	61
18	56	18	54	18	62
19	59	19	57	19	64
20	62	20	60	20	65
21	64	21	63	21	66
22	67	22	66	22	68
23	69	23	69	23	69

SESAM-2A	MP2	SESAM-2B	MP2	SESAM2C	MP2
0	0	0	0	0	0
1	1	1	1	1	2
2	2	2	2	2	5
3	4	3	4	3	8
4	5	4	7	4	10
5	7	5	10	5	13
6	8	6	12	6	17
7	10	7	16	7	20
8	12	8	19	8	24
9	15	9	23	9	27
10	17	10	26	10	30
11	20	11	30	11	34
12	23	12	34	12	37
13	26	13	38	13	40
14	30	14	41	14	43
15	34	15	45	15	46
16	37	16	48	16	48
17	42	17	52	17	51
18	46	18	55	18	54
19	52	19	58	19	56
20	57	20	61	20	59
21	63	21	65	21	61
22	68	22	68	22	63
				23	66
				24	68

SESAM-3A	MP3	SESAM-3B	MP3	SESAM3C	MP3
0	0	0	0	0	0
1	5	1	3	1	3
2	7	2	6	2	6
3	11	3	9	3	10
4	14	4	12	4	14
5	17	5	15	5	18
6	19	6	19	6	22
7	22	7	22	7	25
8	24	8	24	8	29
9	27	9	27	9	32
10	29	10	29	10	34
11	31	11	32	11	37
12	34	12	34	12	40
13	36	13	36	13	42
14	38	14	38	14	45
15	40	15	41	15	48
16	42	16	43	16	50
17	44	17	46	17	53
18	47	18	49	18	57
19	50	19	53	19	60
20	53	20	57	20	64
21	58	21	61	21	67
22	62	22	64		
23	67	23	67		

SESAM-4A	MP4	SESAM-4B	MP4	SESAM4C	MP4
0	0	0	0	0	0
1	2	1	2	1	3
2	5	2	5	2	7
3	7	3	8	3	11
4	9	4	10	4	15
5	12	5	13	5	19
6	14	6	16	6	22
7	17	7	19	7	26
8	20	8	22	8	29
9	23	9	25	9	32
10	26	10	28	10	36
11	29	11	31	11	39
12	32	12	34	12	42
13	36	13	37	13	44
14	40	14	40	14	47
15	43	15	43	15	50
16	47	16	46	16	52
17	51	17	49	17	55
18	55	18	52	18	57
19	58	19	55	19	60
20	62	20	58	20	62
21	66	21	61	21	65
22	69	22	65	22	67
23	72	23	68	23	69
24	75	24	72	24	71
25	77	25	75	25	74
26	80	26	79	26	76
27	82	27	82	27	79
				28	82

SESAM-5A	MP5	SESAM-5B	MP5	SESAM5C	MP5
0	0	0	0	0	0
1	4	1	3	1	2
2	7	2	6	2	4
3	11	3	9	3	7
4	14	4	12	4	10
5	18	5	14	5	13
6	21	6	17	6	16
7	24	7	20	7	20
8	27	8	23	8	23
9	30	9	26	9	27
10	33	10	29	10	30
11	36	11	32	11	33
12	39	12	35	12	37
13	42	13	39	13	40
14	45	14	42	14	44
15	47	15	46	15	47
16	50	16	49	16	51
17	53	17	53	17	54
18	55	18	56	18	57
19	58	19	60	19	60
20	60	20	64	20	63
21	63	21	67	21	65
22	66	22	70	22	68
23	69	23	73	23	71
24	73	24	75	24	73
25	76	25	77	25	76
26	79	26	79	26	78
27	82	27	82	27	80
				28	82

SESAM-6A	MP6	SESAM-6B	MP6	SESAM6C	MP6
0	0	0	0	0	0
1	2	1	6	1	2
2	3	2	10	2	4
3	5	3	14	3	7
4	7	4	17	4	10
5	10	5	20	5	12
6	12	6	23	6	15
7	15	7	25	7	19
8	18	8	28	8	22
9	21	9	31	9	25
10	25	10	33	10	29
11	28	11	36	11	32
12	32	12	38	12	36
13	35	13	41	13	39
14	39	14	43	14	43
15	43	15	46	15	46
16	46	16	48	16	50
17	50	17	51	17	53
18	54	18	53	18	57
19	58	19	56	19	60
20	62	20	58	20	63
21	66	21	61	21	66
22	71	22	64	22	69
23	75	23	67	23	72
24	79	24	70	24	75
25	80	25	73	25	78
26	81	26	76	26	80
27	82	27	79	27	82
		28	82		