

# Implementing content-based mathematics teaching innovations at scale

Ola Helenius <sup>1</sup> 

## Introduction

As researchers, we hope that findings from our field will influence pedagogical practice. One way this can happen is through structured implementation programs. When something is described as an implementation program, it usually implies that a particular innovation is being introduced, and that there is a reasonably well-developed plan for how this innovation is to be enacted within the organization it is intended to affect.

In a review of 104 articles in mathematics education research that self-identified as addressing implementation, we found that only 27 studies focused primarily on specific mathematical content. Other categories instead concerned innovations related to general reform-oriented teaching, particular mathematical competencies or proficiencies, or specific teaching models (Ahl et al., 2023). This tendency toward the general rather than the subject-specific is perhaps unsurprising, given decades of focus on what is often called reform mathematics teaching, alongside a broader shift toward social and sociocultural perspectives on teaching and learning (Inglis & Foster, 2018).

---

<sup>1</sup> ola.helenius@ncm.gu.se

University of Gothenburg, Gothenburg, Sweden

Helenius, O. (2026). Implementing content-based mathematics teaching innovations at scale. In M. S. Aguilar & M. Misfeldt (Eds.), *Proceedings of the First Implementation and Replication Studies in Mathematics Education Conference: IRMEC 2025* (pp. 29–45). Editorial SOMIDEM. <https://doi.org/10.24844/SOMIDEM/S3/2026/03-02>

In this chapter, I use two implementation projects, both involving rather radical changes in how quite basic and specific mathematics is taught, as cases for discussing implementation design. Both projects are relatively large-scale: not national in scope, but large enough that the possibility of direct contact between developers and teachers becomes a central issue in the design of the implementation process.

Since I was invited to write this contribution in a somewhat open format rather than as a conventional research report, I take the opportunity to describe aspects of our design work and experiences that we have not necessarily formalized as research. Indeed, both programs that I will use as cases emerged as research-based development initiatives rather than as research projects in the strict sense. Although they were developed interactively together with practitioners, they are not best characterized as design research projects, as we prioritized relatively rapid development over systematic documentation and analysis of each iteration.

What has been, and is currently being, researched in relation to these projects will be presented as the discussion unfolds. At the same time, I find it necessary to articulate some of the lessons we have learned so far, even when fully formalized studies do not yet support these insights. For this reason, the chapter will not include a separate methods section. Instead, I take a more framework-oriented approach to the discussion, which is introduced in the next section.

In this respect, my approach is in the tradition of Cobb and Jackson (2021) in their discussion of the MIST project, where they describe elements of a coherent instructional system built around shared goals and visions for mathematics teaching and learning within a school system. However, my intention is to keep the implementation aspects discussed here applicable not only to broad and overarching instructional ideas, but also to more minor, content-specific changes in everyday teaching.

### **Theoretical perspectives**

To establish a framework for discussing implementation more structurally, I will first examine implementation ‘from above’ by identifying the components of an implementation effort that can be deliberately planned. I will then complement this perspective with a look ‘from within’, focusing on factors that may influence the success of implementation in practice.

### **Implementation components**

To begin discussing implementation components, it is helpful to start with the issue of scale. Aguilar and colleagues examine scale and scaling through a Delphi-style review and conclude that there is little agreement on what “large-scale” should mean in terms of a specific level of spread. When

distinguishing between small-scale and medium-scale implementation, they identify three characteristic forms of medium-scale scaling, loosely tied to different motives. Scaling may occur in an organic manner within an existing system; it may involve a deliberate effort, for instance, through a multiplier structure; or researchers may seek to scale up in order to examine whether results that hold at a smaller scale remain valid in more diverse contexts (Aguilar et al., 2023).

In a manner similar to Roesken-Winter et al. (2021), Aguilar et al. (2023) characterize differences among scaling variants by highlighting three key aspects: a contact factor, a material factor, and an organizational factor. To avoid confusion with the factors of influence discussed later, I refer to the three factors introduced above as *components* rather than factors. The term “component” is deliberately chosen to suggest a constructionist perspective rather than an analytic one. When you build something, you can think of the process as selecting components and assembling them. What is being metaphorically built here is an implementation program. While discussing the three components already introduced, I will also introduce two closely related ones.

*The contact component* concerns the degree and form of interaction between researchers or developers and innovation users. Many aspects of teaching are notoriously difficult to communicate at a distance. As Clarke and Hollingsworth (2002) argue in their well-known model of teacher change, complex instructional change requires feedback loops between professional development and classroom enactment. It is therefore much easier to support implementation when there is substantial direct contact between developers, who understand the innovation, and teachers, who are expected to enact it. Such collaboration also allows the innovation itself to be adapted to better fit local contexts, as emphasized in the design research paradigm. If developers and teachers are involved in an adaptation process together, it is more likely that core components or determinants of the innovation (Chen, 2012) are preserved. Conversely, when direct contact is limited, the innovation must be communicated through other means, which brings us to the next component.

*The material component* concerns texts, videos, curricular resources, and other artifacts intended to support teachers in understanding and enacting the innovation. Such materials may include complete curricular programs like student textbooks, teacher guides, lesson resources, slide presentations, but may also consist of more theoretical or explanatory texts. Following Kennedy’s (2016) characterization of professional development, it may be helpful to think of the material component as supporting enactment through differing combinations of prescription, strategy, insight, and body of knowledge. Most implementation materials contain elements of all four,

but in different proportions. Materials that attempt to provide a great deal of all four simultaneously may become difficult both to produce and to use.

Before turning to the organizational component, two additional bridging components link developer contact and material support and also function as substructures of the organizational dimension.

*The facilitator component* corresponds to what is sometimes described as a personnel strategy (Roesken-Winter et al., 2021). The facilitator role can vary across several dimensions: facilitators may be more or less expert in the innovation itself, and more or less positioned as organizational leaders. When facilitators act as a bridge between developers and teachers, they are sometimes called multipliers. A facilitator may be a teacher colleague but may also be someone serving in another professional capacity, such as a teacher educator. In some cases, facilitators may already occupy formal roles within the system, similar to what Cobb, Jackson, and colleagues describe as mathematics leaders or district leaders (Jackson et al., 2015). If facilitators take on an expert role, they may develop this expertise through close collaboration with developers or through preparation using implementation materials. In Timperley and colleagues' (2007) review of teacher professional learning, the strategy with the strongest empirical support involved collegial teacher groups working with expert guidance, leading naturally to the next component.

*The collegial group component* concerns teachers working together to interpret, prepare, enact, and evaluate the innovation. Such collaboration can occur with varying frequency and intensity and may take place mainly as preparation before enactment or as recurring meetings closely tied to ongoing enactment. Collegial work may focus on overarching principles, shared goals, and instructional vision, or it may be oriented toward specific lessons or instructional sequences. In some cases, collegial meetings are supported by structured materials that guide discussion toward particular questions or practices.

*The organizational component* is the final component, and it naturally connects the others. While pedagogical leadership may play an important role, implementation design often hinges even more on organizational decisions about time, personnel allocation, and coordination. The organization must support the contact, facilitator, and collegial components, all of which require resources and sustained commitment. Moreover, willingness and capacity to allocate time and personnel may fluctuate. Key individuals may change priorities or be replaced, unless there are more stable systemic structures that support continuity. As demands increase, in terms of personnel, time investment, and coordination across units, the organizational component becomes increasingly critical and more challenging to design.

We began this section with the contact component, referring to direct interaction between teachers and developers. However, as will become important in the next section, there is also a second contact dimension: interaction between developers and organizational leadership at appropriate levels. Designing this interaction is itself a key element of implementation design. I have previously described the setup of local organizational structures as a form of stakeholder scaffolding, and I consider the deliberate establishment of such scaffolding to be part of implementation design (Helenius, 2022).

### **Factors of influence**

In a spin-off from the review mentioned in the introduction (Ahl et al., 2023), Ahl, I, and our colleagues discuss classical categories of implementation, such as those summarized by Century and Cassata (2016): the innovation, the user (in this case, the teacher), the organization, and the external environment. These broad categories alone do not sufficiently capture the factors that shape implementation success. Instead, we propose focusing on the interplay between them (Ahl et al., in press).

First, we identify several influential factors arising at the intersection *between teachers and innovation*. To what extent do teachers perceive a need for the innovation, and do they have the knowledge and confidence required to enact it? To what extent does the innovation interact negatively with existing routines, with parental expectations, or with other established practices? The more demanding the innovation is in relation to these conditions, the more implementation support will be required to manage the teacher-innovation boundary.

Second, we consider factors emerging at the *boundary between the implementation project and the school organization*. As implementations extend over longer periods, involve multiple schools, include collegial collaboration, require teacher-developer contact, and perhaps incorporate multiplier structures, the organizational strain increases. Greater implementation complexity increases the number of potential failure points, thereby generating more influential factors that must be addressed at this intersection. From a design standpoint, added implementation complexity typically implies a stronger need for developers to remain involved in, or at least maintain some form of oversight of, larger parts of the school organization.

A third set of factors can be traced to the *intersection between the innovation and the wider educational environment*. Such factors include external testing regimes, accountability systems, broader pedagogical trends, or externally imposed curricula, all of which may align more or less well with the innovation. There may be good reasons to design innovations that

do not align neatly with testing regimes, national curricula, or similar external factors. However, as with innovations that conflict with teacher routines, such misalignment increases the need for deliberate implementation structures capable of supporting enactment despite external pressures not aligned with program intentions.

### **Two implementation projects: TRR and Algebra Monday**

I will now describe the origins of the two implementation programs, TRR<sup>1</sup> and Algebra Monday, and how the five components from the theory section are embedded in the two programs. In a separate section, I will then compare the programs through a factors-of-influence perspective.

#### **TRR**

TRR was initially established as a development program focused on arithmetic for six-year-olds (grade 0) at the Swedish National Center for Mathematics Education (NCM) around 2011, financed by a national school authority. In this form, it was conducted as a design research project (Sternér & Helenius, 2015) and evaluated in a randomized controlled trial (Sternér et al., 2020). TRR has also been further evaluated (Sternér et al., 2023; Vennberg & Norqvist, 2018). TRR was made into a publicly available teacher guide (Sternér et al., 2014), sold in book form by NCM. In addition to the physical teacher guide, a main product of this TRR-0 incarnation was a Vygotskian-inspired six-phase teaching cycle, typically covering approximately three lessons (Koljonen et al., 2022; Sternér & Helenius, 2015).

At a later stage, the Swedish Association of Local Authorities and Regions (SALAR, but we will use the Swedish acronym SKR), which is an umbrella organization for municipalities, contacted NCM around a professional development program for grade 1–3 teachers. At this point, Sweden was just finishing a national-scale teacher professional development program, the Boost for mathematics, which is now well researched (see Helenius, 2021 for an overview). Preliminary results, later published by Lindvall and colleagues (Lindvall et al., 2018, 2022), were already known and indicated that the Boost for mathematics did not have significant effects on students' learning. Referring to the distinction coined by Grossman and McDonald (2008), instead of aiming for further professional development in the pedagogies of the investigation paradigm, it was decided to explore the pedagogies of the enactment route. We decided to accelerate the already initiated development of TRR-0 towards

---

<sup>1</sup> TRR is an acronym for the Swedish name *Tänka resonera och räkna*, in English: Thinking reasoning and counting.

grades 1–3 and base the collaboration between NCM and SKR on enacting the TRR 1–3 teaching design. The design of the content was based along the lines of Vergnaud's conceptual field theories (Helenius & Ahl, 2025a, 2025b; Säfström et al., 2019; Vergnaud, 1998, 2009) and inspired by work by the likes of Thompson and Saldanha (2003), Mulligan and Mitchelmore (2009), Lamon (1994, 2007), and others (Beckmann & Izsák, 2015; Izsák & Beckmann, 2019; Sowder et al., 1998).

When developing TRR 1–3, we decided to group the six-phase teaching cycles, typically covering three lessons each, into themes with five cycles in each. For every school year, there are four TRR themes. Teachers participating in TRR conduct TRR-based teaching for approximately half of the total teaching time, with two themes in the fall and two in the spring each year.

I will now describe the TRR 1–3 implementation in terms of the implementation components. In this description, I will stick to the first 14 municipalities that implemented TRR 1–3, under the leadership of SKR and NCM. I don't know how many teachers are currently working with TRR, but based on the 14 original municipalities, we once estimated it to around 1000 teachers.

TRR had a pretty weak contact component with only occasional presentations to all teachers who were attending a particular municipality. These presentations were also primarily focused on the intentions and principles underlying TRR rather than on the detailed content. On the other hand, the material component was powerful, with detailed booklets describing every teaching cycle. The booklets also described preferred vocabulary, representations, and artefacts to use and how to use them; exact tasks for each phase of each cycle; how to use students' productions, and so on.

The collegial group component of TRR was arranged by grouping teachers teaching the same grade together, often across schools since many schools only had one or two teachers per grade. Teachers in one group were supposed to teach the same cycle the same week. Before and after every TRR week, these groups met to discuss the TRR cycle for the upcoming week and follow up on the teaching from the previous week. Each theme included a special material component, a small booklet, to guide these meetings. In this sense, the collegial group component was tightly tied to actual teaching practice and to the content of each teaching cycle. The facilitator component of TRR was structured such that each collegial group had a designated leader. The facilitators received a brief introduction from NCM the term before their colleagues began TRR, which also included testing some TRR cycles in their own class. That said, the developer-facilitator contact was still relatively weak, except for a few select so-called

design teachers who tested TRR cycles as they were developed during the years when the TRR implementation was planned.

The organizational component of TRR was powerful. Using expertise and experience from SKR, we designed a specific stakeholder scaffolding, which I have described in detail elsewhere (Helenius, 2022). Each municipality had to assign a process leader and a network group comprising teachers, principals, and administrative and political leadership. In each round of implementation, 4–5 municipalities collaborated with the network groups, which met at biannual network conferences to discuss implementation. Locally, municipalities could decide the scaling, but had to dimension the choice of facilitators beforehand, so that these could be educated by NCM. There also had to be a mandatory local process leader. While contact between NCM (as developer) and the process leaders was relatively frequent, we, as developers, still had limited control over the local implementation.

### **Algebra Monday**

Algebra Monday, unlike TRR, has no organizational backing and comprises only two researchers as developers: me and Linda Marie Ahl (from now on, the AM-developers). Algebra Monday as an implementation project started with the AM-developers getting contacted by two upper secondary teachers who, over many years, have seen declining results and felt an urgent need for some change. As these two teachers were so vital in the early development phase, I will call them AM co-developers. When designing Algebra Monday, my guiding principle was: what would you do if you had to bet \$10,000 of your own money on it actually working, in terms of generating a long-term, broad improvement in student outcomes? At this point, we, as developers, had worked with implementation research for many years and had extensive experience in municipal collaboration. After discussions with municipal leadership, we developed the basic tenets of a comprehensive grade 0 (six-year-olds) to upper secondary (sixteen-year-olds) strategy for algebra and structural arithmetic. A decision based on our implementation research was to have all participating teachers work only once a week on algebra and structured arithmetic, using specific lesson designs with prescribed content developed by us. As we wanted all teachers across all grades to do this consistently, not limited to a certain number of weeks, months, or even years, our aim was extensive from a systems perspective. On the other hand, we did not intend to change any other aspects of the teacher's teaching (except through implicit influence). So even if teachers had to devote one day a week to Algebra Monday, we considered this a much less intrusive design than TRR. We also knew from other design and implementation projects (Teledahl et al., 2023) that a

project with a specific name that was only supposed to take place occasionally, though perhaps rather frequently, was comparatively easy to get acceptance for by both teachers and students.

As it happened, after approximately one year of preparation, the original municipality decided not to proceed with the project. Instead, it was taken up by a neighboring municipality to which the AM co-developers moved and were granted favorable conditions for leading the project locally. I will come back to the details of the organization. During the development phase of Algebra Monday, four other municipalities contacted us and expressed interest in joining, so, in the end, four municipalities started Algebra Monday at the beginning of 2025, with another following in the fall.

Algebra Monday is also entirely focused on the mathematical content. In practice, it also involves a specific lesson design, but this is not foregrounded nearly as much as in TRR. For grades 4 and up, the Algebra Monday design follows a design developed over many years by Ahl, who also coined the name (Ahl & Helenius, 2026). To allow easier comparison with TRR, we will here only discuss the grades 0–3 design, which almost exactly shares the mathematical content with TRR, albeit in a re-worked form with a differently organized progression. Most importantly, though, the Algebra Monday version is designed to fit into single lessons and offer more teacher flexibility than the TRR teaching model.

I will now discuss Algebra Monday from the perspective of the five implementation components from the theory section.

The contact component for Algebra Monday is exceptionally high, particularly for the facilitators/multipliers whom we call pilot teachers, as discussed further below. The typical design is that the developers and pilot teachers meet digitally 1.5–2 hours every week or sometimes every other week, following a specific and very efficient meeting design. At these meetings, the pilot teachers provide feedback to the developers and to one another to gradually improve their enactment of the Algebra Monday lessons. Before the meeting, a PowerPoint presentation specific to each lesson and week is created, and the teacher uploads pictures of students' work, photos of the whiteboard, and perhaps of students in action, sometimes also videos to this common presentation. This arrangement makes it quick and easy to share thoughts from the previous lessons. The content for the next week (or two weeks) is presented, normally by an AM-developer, and thoroughly discussed with the pilot teachers. This is done over 20 effective weeks, that is, until the pilot teachers have delivered 20 Algebra Monday lessons (each with a specific content plan, developed by the AM-developers). We can run such collegial groups with up to 20 pilot teachers in each meeting. The high level of developer-pilot teacher contact

in practice limits Algebra Monday to five municipalities in the current design. The material component for Algebra Monday is relatively modest. Each lesson is presented digitally, and teachers receive a PDF copy of the presentation as a material component. At least initially, it would be impossible to conduct the lesson in any meaningful way using this material without a verbal presentation. The contact component hence powerfully dominates the material component.

The facilitator component has been addressed briefly above. Together with the municipality, we recruit the pilot teachers among the best teachers in the municipality. They should therefore be both self-selected and identified as particularly competent by others. After 20 weeks of Algebra Monday work, the pilot teachers assume the role of facilitators, typically working in pairs to reduce sensitivity. Facilitators then lead collegial groups of new teachers, with one facilitator pair leading perhaps 7–10 other teachers. Typically, the AM-developers assist with collegial work for approximately 10 additional weeks. One type of arrangement is that collegial groups first work for one hour in individual digital rooms with discussions on the previous lesson. Then all groups meet for a presentation of the lesson for the next week, which they then might discuss in separate groups again. Eventually, the facilitators also take responsibility for the presentation of new material. But they will then always present material they themselves used in their own classes before.

The organizational component of Algebra Monday is looser and more diverse than it was for TRR. We will not describe this diversity but rather aim to provide a description that captures the core of the Algebra Monday organization. As it was with the municipality of the AM co-developers, it is always the municipality that contacts us AM-developers and asks if they can join. We AM-developers then make sure that the administrative leadership in the municipality stands behind this decision. We require leadership in the form of process leaders who are themselves teachers and can carry out the Algebra Monday teaching in their classes, but also have some connection to, or authority from, the municipality's administrative leadership. The process leaders get responsibility for all the local organization, including setting up digital meeting rooms, creating appropriate groups, setting up the administration of places to upload documentation to, and so on. What is also required for this setup to work is one single 1.5–2 hour timeslot common to the whole municipality or at least to all participating schools. This is certainly not a standard practice in Sweden, as principals typically organize professional development from a school-based perspective. I will return to the implementation difficulties associated with practical autonomy and the role of principals in the Swedish school system.

### **A comparison of factors of influence between TRR and Algebra Monday**

Let us now discuss TRR and Algebra Monday in terms of the three classes of factors of influence presented in the theory section. I will not try to compare outcomes but rather look at the two implementation programs from a design point of view. When the TRR implementation was designed, we did not have the conceptual understanding of implementation we have today. Yet, I will in a retrospective fashion formulate the design through this theoretical framing. While one of the strengths of Algebra Monday is its grade 0–9 integration, to enable an easier comparison I will as mentioned earlier only discuss the grades 0–3 component.

Regarding the intersection *between teachers and the innovation*, both TRR and Algebra Monday portray the curricular content in a very non-standard manner related to how it is typically done in Sweden (Ahl & Helenius, 2021, 2024; Helenius & Ahl, 2024a, 2024b). This means that the teacher will have to teach number and additive and multiplicative structures in ways that they might never have seen before. The relation between the innovation and teachers' knowledge will hence be a critical factor of influence for both programs. Both programs also build on a teaching design that is non-standard compared to typical Swedish teaching (Boesen et al., 2014). The teaching design is however much more foregrounded in TRR and TRR also covers more time. For Algebra Monday we deliberately limited the time to once a week to make it less intrusive. The lesson design for Algebra Monday is also more flexible.

Regarding the *boundary between the implementation project and the school organization*, the TRR design has a mandatory involvement of principals when municipalities follow the designed stakeholder scaffolding (Helenius, 2022). Principals are obliged to set up times for collegial meetings every week TRR runs. Principals in this sense have a rather important role in TRR. In Algebra Monday, we have deliberately tried to avoid making the program school based. For example, we prefer collegial groups formed across schools and the digital setup makes this possible. There is however a very important demand on principals. All attending teachers, which in some primary schools may be almost all teachers, must be available at the particular time slot decided for this municipality. This might mean some schools have to change a weekly schedule that they have had for years. Even though the municipality administrative leadership technically can force principals to do this, in the Swedish tradition principals have a very high level of autonomy. In this sense Algebra Monday is more intrusive in the school organization than TRR. On the other hand, once the time slot is arranged, Algebra Monday relies less on organizational school issues than TRR. Algebra Monday has a higher demand on the process leaders than TRR though, since the process leaders are in charge of arranging the

meeting structure. The last category of factors of influence concerns the *intersection between the innovation and the wider educational environment*. During the years when both TRR and Algebra Monday were started, the Swedish national curriculum had a very low level of specification related to the content treated by TRR and Algebra Monday (Helenius & Ahl, 2024b; Prytz, 2023a, 2023b). There seems to be a trend towards a more algebraic and structural way of treating arithmetic for mathematics textbooks for the primary grades though. There is also ongoing work in Sweden towards a new national curriculum where the early algebra and structural arithmetic approach advocated in both TRR and Algebra Monday is strengthened (Helenius, 2025) which will mean that both Algebra Monday and TRR become better aligned with the wider educational environment.

## Discussion

By mapping TRR 1–3 and the grades 0–3 part of Algebra Monday in terms of the implementation components introduced in the theory section, I have shown how the two programs come out as very different from an implementation perspective, despite building on almost identical content, underlying theory, and ideas around the design of teaching. The two programs also share the same focus on the actual teaching practice, making them both fit squarely in the pedagogies of enactment paradigm (Grossman & McDonald, 2008), which puts them in the prescription-strategy end of the PD-spectrum rather than in the insight-body of knowledge end of Kennedy's spectrum (Kennedy, 2016). I propose that mapping the five implementation components of an implementation program can be useful when comparing different implementation efforts. While the five implementation components do not offer the same level of detail as for example the structured reporting guidelines for interventions suggested by Hill and colleagues (Hill et al., 2024), the five components I have suggested might be more useful from a design point of view. As a designer of an implementation, your options for configuring the different components are often constrained by surrounding factors or early choices. For example, if aiming for scaling in terms of a large spread (Coburn, 2003), a strong contact component is typically impossible, so you are left to choose between a strong material component like in TRR, or a complex multiplier structure, perhaps arranged in several layers with level-1 multipliers being educated by the developers and level-2 multipliers being educated by level-1 multipliers.

Making such a complex design work will however need a very strong connection with the school organization and the system it sits in, which Prediger and Selter have discussed in terms of a shared vision on all levels (Prediger & Selter, 2024).

Moving on to the factors of influence analysis I also carried out with TRR and Algebra Monday as cases, such an analysis should in my opinion, be carried out as a part of the design process, not only be something that guides evaluation. In particular, the intersection between teachers and the innovation should be carefully analyzed before designing the rest of the implementation. The bigger the discrepancy between the requirements of the innovation and teacher knowledge and existing routines, the stronger the support structure built into the implementation needs to be. A strong support factor in terms of time, collegial arrangements, material support and developer contact will in turn make the organization more complex which in itself introduces more possible failure points.

## Acknowledgments

The following grants financed this work: Swedish Research Council Grant (2023–2026), Marcus and Amalia Wallenberg Foundation Grant (2025–2027), and Swedish Institute for Educational Research Grant (202400048).

## References

- Aguilar, M. S., Ahl, L. M., Jankvist, U. T., & Helenius, O. (2023). Towards characterization of scale and scaling in implementation research within mathematics education. *Implementation and Replication Studies in Mathematics Education*, 3(1), 99–122. <https://doi.org/10.1163/26670127-bja10012>
- Ahl, L. M., & Helenius, O. (2021). A framework for analyzing progress in concept knowledge in mathematics textbooks. In G. A. Nortvedt, N. A. Buchholtz, J. Fauskanger, F. Hreinsdóttir, M. Häikiöniemi, B. E. Jessen, J. Kurvits, Y. Liljekvist, M. Misfeldt, M. Naalsund, H. K. Nilsen, G. Pálsdóttir, P. Portaankorva-Koivisto, J. Radišić, & A. Wernberg (Eds.), *Bringing Nordic mathematics education into the future. Proceedings of Norma 20. The ninth Nordic Conference on Mathematics Education Oslo, 2021* (pp. 1–8). SMDF.
- Ahl, L. M., & Helenius, O. (2024). Is it a fraction or should I divide it? In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.), *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 17–24). IGPME.
- Ahl, L. M., & Helenius, O. (2026). Designing and scaling theory-informed algebra instruction in Sweden. In M. S. Aguilar & M. Misfeldt (Eds.), *Proceedings of the First Implementation and Replication Studies in Mathematics Education Conference: IRMEC 2025* (pp. 129–132). Editorial SOMIDEM. <https://doi.org/10.24844/SOMIDEM/S3/2026/03-16>
- Ahl, L. M., Helenius, O., Aguilar, M. S., Jankvist, U. T., Misfeldt, M., & Prytz, J. (2023). Implementation research in mathematics education: A systematic mapping review. *Implementation and Replication Studies in Mathematics Education*, 3(2), 135–199. <https://doi.org/10.1163/26670127-bja10015>

- Ahl, L. M., Helenius, O., Jankvist, U. T., Aguilar, M. S., & Prytz, J. (in press). From local reform to systemic change: A three-dimensional model for large-scale implementation in mathematics education. *Implementation and Replication Studies in Mathematics Education*.
- Beckmann, S., & Izsák, A. (2015). Two perspectives on proportional relationships: Extending complementary origins of multiplication in terms of quantities. *Journal for Research in Mathematics Education*, 46(1), 17–38. <https://doi.org/10.5951/jresmetheduc.46.1.0017>
- Boesen, J., Helenius, O., Bergqvist, E., Bergqvist, T., Lithner, J., Palm, T., & Palmberg, B. (2014). Developing mathematical competence: From the intended to the enacted curriculum. *The Journal of Mathematical Behavior*, 33, 72–87. <https://doi.org/10.1016/j.jmathb.2013.10.001>
- Century, J., & Cassata, A. (2016). Implementation research: Finding common ground on what, how, why, where, and who. *Review of Research in Education*, 40(1), 169–215. <https://doi.org/10.3102/0091732X16665332>
- Chen, H. -T. (2012). Theory-driven evaluation: Conceptual framework, application and advancement. In R. Strobl, O. Lobermeier, & W. Heitmeyer (Eds.), *Evaluation von Programmen und Projekten für eine demokratische Kultur* (pp. 17–40). Springer. [https://doi.org/10.1007/978-3-531-19009-9\\_2](https://doi.org/10.1007/978-3-531-19009-9_2)
- Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, 18(8), 947–967. [https://doi.org/10.1016/S0742-051X\(02\)00053-7](https://doi.org/10.1016/S0742-051X(02)00053-7)
- Cobb, P., & Jackson, K. (2021). An empirically grounded system of supports for improving the quality of mathematics teaching on a large scale. *Implementation and Replication Studies in Mathematics Education*, 1(1), 77–110. <https://doi.org/10.1163/26670127-01010004>
- Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, 32(6), 3–12. <https://doi.org/10.3102/0013189X032006003>
- Grossman, P., & McDonald, M. (2008). Back to the future: Directions for research in teaching and teacher education. *American Educational Research Journal*, 45(1), 184–205. <https://doi.org/10.3102/0002831207312906>
- Helenius, O. (2021). A stakeholder analysis of the development of a large scale professional development program. *Implementation and Replication Studies in Mathematics Education*, 1(2), 227–256. <https://doi.org/10.1163/26670127-01010009>
- Helenius, O. (2022). Designing a stakeholder scaffolding for implementing a teaching model at scale. In J. Hodgen, E. Geraniou, G. Bolondi, & F. Ferretti (Eds.), *Proceedings of the Twelfth Congress of the European Society for Research in Mathematics Education (CERME12)* (pp. 4145–4152). Free University of Bozen-Bolzano and ERME. <https://hal.science/hal-03765106/document>

- Helenius, O. (2025). *Ämnesspecifika instruktioner för ämnet matematik*. Göteborgs universitet. <https://www.skolverket.se/download/18.6d479e9619a0a1e68431c9a/1761752287393/Ämnesspecifika%20instruktioner%20-%20Matematik.pdf>
- Helenius, O., & Ahl, L. M. (2024a). A framework for analyzing long-term early algebra progression in textbook series. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.), *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 9–16). IGPME.
- Helenius, O., & Ahl, L. M. (2024b). *Läromedel på villovägar – en analys av den bristande vetenskapliga grunden för dagens svenska läromedel i matematik* (Näringslivets skolforum). Svenskt Näringsliv.
- Helenius, O., & Ahl, L. M. (2025a). Conceptual fields and conceptual polysemy of mathematical objects with applications to the multiplicative conceptual field. *Hiroshima Journal of Mathematics Education*, 18, 51–69. <https://doi.org/10.24529/hjme.1804>
- Helenius, O., & Ahl, L. M. (2025b). Gérard Vergnaud in action. *For the Learning of Mathematics*, 45(1), 50–55.
- Hill, H. C., Lynch, K., & Polanin, J. R. (2024). *Structured reporting guidelines for classroom intervention research* (EdWorkingPaper No. 24-1022). Annenberg Institute for School Reform at Brown University. <https://eric.ed.gov/?id=ED660203>
- Inglis, M., & Foster, C. (2018). Five decades of mathematics education research. *Journal for Research in Mathematics Education*, 49(4), 462–500. <https://doi.org/10.5951/jresmetheduc.49.4.0462>
- Izsák, A., & Beckmann, S. (2019). Developing a coherent approach to multiplication and measurement. *Educational Studies in Mathematics*, 101(1), 83–103. <https://doi.org/10.1007/s10649-019-09885-8>
- Jackson, K., Cobb, P., Wilson, J., Webster, M., Dunlap, C., & Appelgate, M. (2015). Investigating the development of mathematics leaders' capacity to support teachers' learning on a large scale. *ZDM – Mathematics Education*, 47(1), 93–104. <https://doi.org/10.1007/s11858-014-0652-5>
- Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, 86(4), 945–980. <https://doi.org/10.3102/0034654315626800>
- Koljonen, T., Helenius, O., & Ahl, L. M. (2022). The documental approach of didactics as a gauge of fidelity of a teaching model implementation. In J. Hodgen, E. Geraniou, G. Bolondi, & F. Ferretti (Eds.), *Proceedings of the Twelfth Congress of the European Society for Research in Mathematics Education (CERME12)* (pp. 4161–4168). Free University of Bozen-Bolzano and ERME.

- Lamon, S. J. (1994). Ratio and proportion: Cognitive foundations in unitizing and norming. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 89–120). State University of New York Press.
- Lamon, S. J. (2007). Rational numbers and proportional reasoning: Toward a theoretical framework for research. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 1, pp. 629–667). Information Age Publishing.
- Lindvall, J., Helenius, O., Eriksson, K., & Ryve, A. (2022). Impact and design of a national-scale professional development program for mathematics teachers. *Scandinavian Journal of Educational Research*, 66(5), 744–759. <https://doi.org/10.1080/00313831.2021.1910563>
- Lindvall, J., Helenius, O., & Wiberg, M. (2018). Critical features of professional development programs: Comparing content focus and impact of two large-scale programs. *Teaching and Teacher Education*, 70, 121–131. <https://doi.org/10.1016/j.tate.2017.11.013>
- Mulligan, J., & Mitchelmore, M. (2009). Awareness of pattern and structure in early mathematical development. *Mathematics Education Research Journal*, 21(2), 33–49. <https://doi.org/10.1007/bf03217544>
- Prediger, S., & Selter, C. (2024). Establish shared visions and support productive adaptations on all levels: Aims, strategies, and architecture of a nationwide implementation program. *Implementation and Replication Studies in Mathematics Education*, 4(1), 15–49. <https://doi.org/10.1163/26670127-bja10020>
- Prytz, J. (2023a). *En ny kursplan i matematik – hur den skulle kunna se ut och varför*. Svenskt Näringsliv.
- Prytz, J. (2023b). *Grundskolans kursplaner i matematik – igår, idag och imorgon*. Svenskt Näringsliv.
- Roesken-Winter, B., Stahnke, R., Prediger, S., & Gasteiger, H. (2021). Towards a research base for implementation strategies addressing mathematics teachers and facilitators. *ZDM – Mathematics Education*, 53(5), 1007–1019. <https://doi.org/10.1007/s11858-021-01220-x>
- Säfström, A., Helenius, O., & Ahl, L. M. (2019). Implementing alternative models for introducing multiplication. In U. T. Jankvist, M. van den Heuvel-Panhuizen, & M. Veldhuis (Eds.), *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education* (pp. 4447–4454). Freudenthal Group; Freudenthal Institute; Utrecht University; ERME.
- Sowder, J., Armstrong, B., Lamon, S., Simon, M., Sowder, L., & Thompson, A. (1998). Educating teachers to teach multiplicative structures in the middle grades. *Journal of Mathematics Teacher Education*, 1(2), 127–155. <https://doi.org/10.1023/a:1009980419975>

- Sterner, G., & Helenius, O. (2015). Number by reasoning and representations – the design and theory of an intervention program for preschool class in Sweden. In O. Helenius, A. Engström, T. Meaney, P. Nilsson, E. Norén, J. Sayers, & M. Österholm (Eds.), *Development of mathematics teaching: Design, scale, effects* (pp. 159–168). Svensk förening för matematikdidaktisk forskning (SMDF).
- Sterner, G., Helenius, O., & Wallby, K. (2014). *Tänka, resonera och räkna i förskoleklassen*. Nationellt Centrum för Matematikutbildning (NCM).
- Sterner, G., Nagy, C., & Nyström, P. (2023). A scaled-up mathematics intervention in preschool classes. *Scandinavian Journal of Educational Research*, 68(7), 1371–1381. <https://doi.org/10.1080/00313831.2023.2250352>
- Sterner, G., Wolff, U., & Helenius, O. (2020). Reasoning about representations: Effects of an early math intervention. *Scandinavian Journal of Educational Research*, 64(5), 782–800. <https://doi.org/10.1080/00313831.2019.1600579>
- Teledahl, A., Kilhamn, C., Helenius, O., & Ahl, L. M. (2023). Writing down the solution separate from solving the problem. *For the Learning of Mathematics*, 43(2), 43–49.
- Thompson, P. W., & Saldanha, L. A. (2003). Fractions and multiplicative reasoning. In J. Kilpatrick, G. W. Martin, & D. Schifter (Eds.), *A research companion to Principles and standards for school mathematics* (pp. 95–113). NCTM.
- Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development: Best evidence synthesis iteration*. Ministry of Education.
- Vennberg, H., & Norqvist, M. (2018). Counting on: Long term effects of an early intervention programme. In E. Bergqvist, M. Österholm, C. Granberg, & L. Sumpter (Eds.), *Proceedings of the 42nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 355–362). PME.
- Vergnaud, G. (1998). A comprehensive theory of representation for mathematics education. *The Journal of Mathematical Behavior*, 17(2), 167–181. [https://doi.org/10.1016/S0364-0213\(99\)80057-3](https://doi.org/10.1016/S0364-0213(99)80057-3)
- Vergnaud, G. (2009). The theory of conceptual fields. *Human Development*, 52(2), 83–94. <https://doi.org/10.1159/000202727>

