

An introduction to TWG23: Implementation of research findings in mathematics education

Linda Marie Ahl¹, Ola Helenius¹, Boris Koichu², Julie Vangsøe Færch³ and Mario Sánchez Aguilar⁴

¹University of Gothenburg, Sweden; linda.ahl@ncm.gu.se

²Weizmann Institute of Science, Department of Science Teaching, Israel

³University College Copenhagen, Denmark

⁴Instituto Politécnico Nacional, CICATA Legaria, Mexico

Introduction

This introduction outlines the thirteen contributions presented at TWG23 during CERME14. The content of the contributions revolves around seven pivotal themes, to be elaborated on below. Furthermore, three thematic discussions took place, emphasizing the following key points: Core Elements of Innovation and Fidelity of Implementation; The Life Cycle of Implementation and Criteria for Success; and Conflicting Goals, Perspectives, and Theories in Implementation. The thematic discussions in TWG23 at CERME14 were not explicitly related to any particular presentation but instead concerned a holistic view of the implementation research field.

Contributions to the TWG23 at CERME14

The TWG23 at CERME14 was led by Ola Helenius, Boris Koichu, Linda Marie Ahl, and Julie Vangsøe Færch. The TWG23 accepted 11 papers and 2 posters, addressing different aspects of implementation research in mathematics education. The authors of the contributions came from Denmark, Israel, Sweden, the Faroe Islands, and the United Kingdom. The contributions were organized into thematic categories:

1. Reporting guidelines for implementation research
2. Design principles
3. Fidelity and adaptations
4. Scaling issues
5. Replication
6. The role of Habits
7. Pre-, present-, and post-implementation issues

Reporting guidelines for implementation research

The session began with a poster by David Nordqvist, Tina Åkegård, Ola Helenius, and Linda Marie Ahl, exploring what should be reported in implementation research. Hill, Lynch, and Polanin (2024) recently highlighted that inconsistent reporting of classroom interventions impedes knowledge accumulation. They recommend structured guidelines that detail aspects such as intervention content, teacher preparation, and implementation support. Nordqvist and colleagues tested this framework by applying it to "Algebra Monday," a program aiming to strengthen foundational algebra across grades. Using a cascading scaling model, the program trains lead teachers, who in turn train others within their municipality. The researchers found Hill et al.'s guidelines to be broadly applicable, although they lacked elements specific to implementation and scaling. They concluded that the framework is promising but needs to be extended to

capture scaling design processes.

Design principles

Julia Tsygan presented a study on implementing the MUSK model to enhance students' mathematical communication. Instead of the typical fidelity-versus-adaptation lens, the study used an integrity framework, examining how core elements were preserved while adapting to local needs. Using the Professional Obligations model (Herbst & Chazan, 2012), it explored how institutional and practical demands shaped decisions. Adaptations, such as feedback cycles and explicit reasoning instruction, made the model more sensible, although they also revealed tensions between innovation design and school conditions. The study emphasized the need for alignment with professional norms and assessment structures.

In the Faroe Islands, a project used design research to address declining PISA performance. With support from the Ministry of Education, the initiative applies iterative, classroom-based principles to bridge theory and practice. Building on the work of Century and Cassata (2016) and Koichu et al. (2021), this framework views innovation as an ecological disruption and promotes teacher collaboration through Communities of Practice. Project MAC targets mathematical modeling and digital tool use through three guiding principles: problem formulation, digital tool integration, and teacher collaboration. It operates across three schools, emphasizing iterative cycles and combined evaluations to inform broader application.

Fidelity and adaptations

Hanna Palmér, Jessica Elofsson, and Camilla Björklund explored fidelity and adaptation in scaling a preschool math intervention. Using frameworks from Century and Cassata (2016) and Coburn (2003), they evaluated how key elements were preserved or adapted. Coburn's dimensions—depth, sustainability, spread, and ownership—guided the analysis. Though early results were promising, scaling raised challenges in maintaining integrity. Productive adaptations occurred when teachers understood and applied core concepts in a flexible manner. Rigid use or superficial changes often limited effectiveness. The authors noted that local contexts, like Sweden's emphasis on preschool teacher autonomy, significantly influence outcomes.

Mark Boylan's contribution examined fidelity through two lenses: design and theory. In mathematics education, where teacher agency is vital, striking a balance between these is essential for effective scaling. Using the Maths Excellence programs as a case, Boylan showed how different components may require different fidelity approaches, aligning core principles with contextual flexibility.

Scaling issues

Uffe Thomas Jankvist, Linda Marie Ahl, Johan Prytz, and Iresha Ratnayake presented a review showing that implementation success varies by scale. Small-scale projects benefit from direct links between developers and practitioners. Medium-scale efforts rely on structured support and voluntary participation. Large-scale initiatives often succeed when focusing on curriculum rather than pedagogy, as altering teaching practices proves to be more challenging. Scaling pedagogical change requires deep, context-specific support rather than top-down mandates.

Tomas Højgaard discussed the Danish Ishøj Project, which aimed to embed competency-based teaching through school-based seminars and teacher collaboration. While the initiative reached all K-9 math

teachers in the municipality and showed progress, it faced challenges such as limited supervisor expertise, leadership gaps, and constrained resources. Effective scaling, the study concludes, needs strong support systems and leadership commitment.

The MIST (Middle-School Mathematics and the Institutional Setting of Teaching) project, presented by Johan Prytz, Iresha Ratnayake, Linda Marie Ahl, and Uffe Thomas Jankvist, used a Design-Based Implementation Research approach. MIST's Theory of Action focused on leveraging district leadership to enhance teaching quality. The project emphasized structured communication and accountability, with district leaders translating feedback into professional development strategies. MIST's success depended on systemic alignment, suggesting that replicating it requires similar leadership capacity and infrastructure.

Julie Vangsøe Færch presented the only replication study, which examined young students' proportional reasoning. The discussion centered on the role of replication in validating findings, understanding contextual limits, and exploring how adaptations affect outcomes. The overlap between replication and implementation research was acknowledged, with both focusing on variability and context.

Habits and Routines

Andrew Grant explored how interleaved practice challenges teachers' planning habits. Although not unique to mathematics education, this practice necessitates adjustments in how teachers structure lessons. For successful implementation, habit formation strategies must be integrated into professional routines, ensuring that innovations become an integral part of daily teaching.

Pre-, Present-, and Post-implementation projects

Morten Misfeldt and colleagues investigated the integration of programming into math education, which remains a challenge. Tools like Scratch can introduce unfamiliar content that clashes with standard curricula. The team distinguished between Computational Thinking (CT), which aligns with existing structures, and Computational Literacy (CL), which requires broader pedagogical change. Based on Piagetian theory, CT is seen as assimilation, while CL reflects accommodation. The study illustrated how each reshapes geometry learning differently.

Jason Cooper and Boris Koichu discussed post-implementation sustainability. In Israel, some teachers continued to engage in problem-solving after formal incentives ended, motivated by enjoyment. Although classroom integration remained limited, their ongoing involvement suggested increased agency and readiness for future innovation.

Finally, Cecilie Carlsen Bach and colleagues introduced a poster on Technology Comprehension (TC) in math education. Using speculative tarot cards, they encouraged teachers to imagine future digital learning scenarios. This participatory approach aimed to foster ownership and align educational innovations with the perspectives of both teachers and students, countering top-down impositions.

Emerging Themes in Educational Implementation Research

One strength of the CERME conference, particularly within smaller discussion groups like TWG23, is the time available for deep dialogue. This summary captures overarching discussions not tied to individual papers but relevant to the broader field of implementation research.

Core Elements of Innovation and Fidelity of Implementation

There is increasing agreement that innovations should clearly define their core components. These "non-negotiables" support fidelity but must still allow space for context-sensitive adaptations. Determining how specific these elements should be remains a challenge, especially as implementation moves from research to practice. Definitions of *productive adaptation* differ, and while overarching concepts are applicable, local adaptations vary widely. Teachers often carry the burden of interpreting innovations, so involving them in co-design from the outset was identified as key to balancing clarity and flexibility.

The Life Cycle of Implementation and Criteria for Success

Implementation efforts typically unfold across dynamic life cycles influenced by institutional backing, funding, and stakeholder engagement. Our discussions raised questions about how to define success, whether in terms of short-term project outcomes or longer-term, systemic change. Readiness for change, both at the individual and institutional levels, was considered essential but challenging to sustain over time. Success criteria should include fidelity of implementation, stakeholder learning, and sustainability, in addition to student learning outcomes.

Conflicting Goals, Perspectives, and Theories in Implementation

The group acknowledged that differing goals and perspectives often shape implementation efforts. At the practitioner level, clarity on shared strategies or theories of change is usually lacking. Similarly, within the research community, debates continue about how much conceptual agreement is necessary to advance the field. This diversity can be productive, but may also create barriers for newcomers. A significant theme was the inherently political nature of educational settings, in contrast to more technical fields. Ignoring these dynamics risks oversimplifying implementation. Researchers were encouraged to address policy, power, and ideology explicitly, as exemplified in Mark Boylan's work.

Future Themes and Research Directions

The group identified key future directions that may be explored further at CERME15:

- Specifying core elements across phases and contexts
- Defining and evaluating success and failure
- Understanding transitions between small-, medium-, and large-scale implementations
- Revisiting foundational literature to develop updated theoretical tools

The discussions reinforced CERME's spirit of collaboration and thoughtful critique. Building implementation research within mathematics education requires shared commitment, open dialogue, and a culture that supports both rigor and inclusivity.