SURVEY PAPER

Implementation-related research in mathematics education: the search for identity

Boris Koichu[1] · Mario Sánchez Aguilar[2] · Morten Misfeldt[3]

Abstract

Implementation has always been a paramount concern of mathematics education, but only recently has the conceptualizing and theorizing work on implementation as a phenomenon begun in our field. In this survey paper, we conduct a hermeneutic review of mathematics education research identified as related to the implementation problematics. The first cycle of the review is based on examples of studies published in mathematics education journals during the last 40 years. It is organized according to five reasons for developing implementation research. The second cycle concerns 15 papers included in this special issue and is organized by four themes, as follows: objects of implementation, stakeholders in implementation, implementation vs. scaling up, and implementability of mathematics education research. The paper is concluded with a refined glossary of implementation-related terms and suggestions for future research.

Keywords Implementation · Implementation research · Innovation · Implementability of mathematics education research · Scaling up · Stakeholders in implementation

1 Introduction

As long as research and practice co-exist in the field of mathematics education, implementation and implementability have always been of paramount concern. Already in 1979, Bruckheimer pointed out the difficulty of implementing innovative curriculum materials in the mathematics classroom, and indicated the existence of "inhibitors of curriculum implementation" (Bruckheimer, 1979, p. 44). He argued that these inhibitors can be properties of the specific curriculum design, but they can also result from an interplay with the overall structure of the educational system. In 1996, Niss identified the *implementation problem* as one of the main problems of mathematics education research that has to do with establishing the structural and organizational framework for conducting mathematics education, providing the necessary resources for conducting mathematics teaching, and addressing issues related to the philosophy and modes of assessment (Niss, 1996). In his characterization of mathematics education as an academic discipline, Niss (1999) suggested that the implementation problem calls for theoretical scrutiny:

> [i]t is fair to claim that the overarching, ultimate end of the whole enterprise is to promote/improve students' learning of mathematics and acquisition of mathematical competencies. It is worth pointing out that the very specification of the terms just used ('promote', 'improve' 'students' (what students are being considered?) 'learning', 'mathematics', 'acquisition', 'mathematical competencies') is in itself a genuine didactic task. (p. 5)

Niss identified the need to theorize such terms as 'promote' and 'improve' that to a large extent express the *raison d'être* for implementation of mathematics education research in mathematics education practice. However, it is only recently that conceptualizing and theorizing work on implementation has begun in our field, though it has been pursued for a while in some other fields, most notably in health care (e.g., Eccles & Mittman, 2006).

Let us recall several landmarks. Confrey et al. (2000) introduced the idea of 'implementation research' as a

✉ Boris Koichu
boris.koichu@weizmann.ac.il

[1] Weizmann Institute of Science, Rehovot, Israel

[2] CICATA Legaria, Instituto Politécnico Nacional, Mexico City, Mexico

[3] University of Copenhagen, Copenhagen, Denmark

means to link applied psychology and systemic reform in mathematics education. Burkhardt and Schoenfeld (2003) reflected on implementation in mathematics education from the perspective of bridging the gap between research and practice by developing appropriate tools and structures. Remillard (2005) suggested that a particular case of implementation—'curriculum use'—can be usefully explored as a participatory relationship between teachers and curriculum materials (this idea was further developed in research on curriculum ergonomics, Choppin et al., 2018). Maaß and Artigue (2013) interpreted implementation as setting a planned intervention or innovation in motion so that new research results lead to the development of new interventions and these are further disseminated. Cai et al. (2017) considered 'classroom implementation', that is, implementation of research-based learning opportunities in the classroom. All these approaches are related, but put forwards different aspects of the implementation problem and imply complementary agendas for action.

The diversity of approaches to implementation was particularly salient in two collections of papers written by participants in Thematic Working Group 23 of the 10th and 11th Congress of the European Society for Research in Mathematics Education (CERME). This group was established in 2017 at CERME10 with a special focus on implementation and replication research (Jankvist et al., 2017). Participants of CERME11 collaboratively produced a tentative conceptualization of implementation in mathematics education in an attempt to accommodate the diversity (see Aguilar et al., 2019, p. 8). A slight modification of that definition was used in the call for papers for this special issue. In this call, *implementation* was referred to as a change-oriented process of endorsing an action plan based on a relatively well-defined resource (such as a research finding, a digital tool, a curriculum, a textbook, or an institutional policy) that occurs in interaction of a community of the resource proponents and a community of the resource adapters, leading to a gradual shift in agency over the resource and the action plan, from the proponents to the adapters, and also leading to changes in communication and practice of both communities. (This definition is further refined towards the end of this paper).

Even more recently, Artigue (2021) analyzed theoretical resources that are likely to support implementation studies, and argued that resources either internal or external to the field of mathematics education can be useful, though in different ways. In addition, Cobb and Jackson (2021) presented elements of *theory of action* as an empirically-grounded theoretical resource for guiding implementation initiatives at scale. Of special note is that Artigue's and Cobb and Jackson's (2021) papers were published in an inauguration issue of a new journal entitled *Implementation and Replication Studies in Mathematics Education* (IRME). The appearance of IRME is another indication of how vibrant and timely the implementation and implementability problematique in mathematics education is today.

This special issue (hereafter SI) aims at further foregrounding the implementation problem, by showing different ways to conceptualize and work with the implementation aspects of mathematics education research. The motivation for this focus is the observation that implementation still more often than not remains in the background, and not in the forefront of mathematics education research. Fifteen papers of this SI essentially represent the specificity of implementation and implementation-related research in mathematics education. In line with Artigue (2021), we presume that though implementation in mathematics education can be informed by theoretical developments in other fields of study (e.g., Century & Cassata, 2016; Nilsen, 2015), the specificity of the mathematics education ecology (Blomhøj, 2021) requires considering implementation in connection with epistemology and ontology of mathematics as a discipline/subject, as well as in relation to the specific ways in which mathematics education is organized nationally and internationally.

In order to better understand the nature of implementation in mathematics education as it has been developing so far, we first survey past empirical research concerning aspects of implementation in mathematics education without theorizing implementation as a phenomenon (Sect. 3). We then provide a detailed introduction to all the papers of this SI (Sect. 4). A characteristic feature of this SI is that the authors—an international group of scholars who were invited to share their work designed a-priori as 'implementation research' or present their 'regular' research via implementation-related lenses—are very explicit about what implementation and implementation research means for them. Since we are entering a research realm that is not yet fully institutionalized and agreed upon, we choose to conduct the survey in accordance with a hermeneutic approach to reviewing the literature; the details of this approach are explained in Sect. 2. We conclude, in Sect. 5, with a refined glossary of implementation-related terms and suggestions for future research.

2 Hermeneutic approach to the literature on implementation in mathematics education

Although hermeneutic literature reviews are used in such fields as health science and well-being (Lawler et al., 2019; Valentine et al., 2021), they are not common in educational research. According to Boell and Cecez-Kecmanovic (2014), a hermeneutic literature review presumes continuous engagement with a growing body of literature, during

which increased understanding and insights are sought. This approach allows engaging in a dialogue with the literature in search of new meanings (Smythe & Spence, 2012). Practically, a hermeneutic literature review is guided by two circles, namely, search and acquisition, and analysis and interpretation. In the first circle, publications relevant to the topic of interest are identified. Here one often prioritizes working in depth with a relatively small set of highly relevant publications, and allows the reading to influence the search and inclusion criteria in an iterative fashion (Boell & Cecez-Kecmanovic, 2014). The analysis and interpretation circle is developed through the reading of the selected publications, and leads to classifications, critical assessment, and development of arguments about the studied topic.

Accordingly, the aim of the hermeneutic review is not an exhaustive characterization of the chosen topic, but rather to reach a saturation point, whereby insights from an in-depth reading of part of the body of the literature are sufficiently comprehensive. This point is evident from the diminished feeling of novelty when reading additional literature. However, this saturation is where one of the limitations of a hermeneutic literature review lies: it has the goal of obtaining insights without considering all evidence surrounding a particular issue as in a systematic literature review. In the case of the hermeneutical review that we report here, it is very likely that certain literature that might be considered under the umbrella of 'implementation research' has not been considered, due to the design of the review itself, for example, by the choice of the initial inclusion criteria.

In our case, we began the search by exploring the research journals in the field of mathematics education, according to Williams and Leatham's (2017) classifications.[1] By means of the review management software *Covidence* (http://www.covidence.org), we identified 137 papers published in these journals during the last 40 years, which contain the key term 'implement' (or derived terms such as 'implementation' or 'implementing') in the title, in the abstract, or in the keywords. Browsing through the located papers showed a huge variety of ways in which the word 'implementation' was used. This led us to introduce an additional inclusion criterion: we decided that the papers of interest should not only concern possible implementation of the presented findings, but conform to the working definition of *implementation*

research proposed in Century and Cassata's (2016) overview of the landscape of implementation research in education. In particular, they provided a conceptual ground made up of notions and distinctions, including their definitions of innovation, implementation and implementation research. These definitions were often cited by the participants in the TSG23 of CERME11. According to Century and Cassata (2016), *implementation research* is "a systematic inquiry regarding innovations enacted in controlled settings or in ordinary practice, the factors that influence innovation enactment, and the relationships between innovations, influential factors, and outcomes." (p. 170).

Thus, among the papers identified, we looked for those papers that empirically addressed a research question on the implementation of an innovation, probably among other research questions. These papers were then grouped according to five reasons for developing implementation research suggested by Century and Cassata (2016). The reasons are as follows: (i) inform innovation design and development; (ii) understand whether (and to what extent) the innovation achieves desired outcomes for the target population; (iii) understand relationships between influential factors, innovation enactment, and outcomes; (iv) improve innovation design, use, and support in practice settings; (v) develop theory (ibid, p. 174). The interpretation of selected papers according to the five reasons is presented in Sect. 3. Within each reason-category, the interpretation went beyond the motivation for conducting the research, and attended to how it was conducted and what was discovered or concluded. This extended focus eventually led us to the development of four new themes, as follows: (i) objects of implementation, (ii) configurations of stakeholders in implementation; (iii) implementation vs. scaling up; (iv) the question of implementability of mathematics education research. These themes emerge in Sect. 3 with regard to past research, and are then systematically put forth in Sect. 4 in relation to the papers included in this SI. That is, whereas Century and Cassata's (2016) five reasons were chosen at the early stage of our reading as a tool to frame and demarcate the past literature, the four themes emerged from this reading and were further used as a frame for discussing papers included in this SI. Eventually, these four themes served as a springboard for a discussion of what implementation-related research in mathematics education currently is, and of how this SI contributes to the field.

---

[1] We refer to Educational Studies in Mathematics, Journal for Research in Mathematics Education, The Journal of Mathematical Behavior, For the Learning of Mathematics, Mathematical Thinking and Learning, Journal of Mathematics Teacher Education, ZDM—Mathematics Education, Mathematics Education Research Journal, International Journal of Mathematical Education in Science and Technology, School Science and Mathematics, International Journal of Science and Mathematics Education, Recherches en Didactique des Mathématiques, and Research in Mathematics Education.

3 Implementation research in the mathematics education literature

As explained, this survey embraces studies that are not explicitly identified by their authors as 'implementation research', but are interpreted as such by us, in accordance

with the working definition offered by Century and Cassata (2016). We use Century and Cassata's (2016) five reasons as an organizational structure, and illustrate each reason through a small number of studies. Each study was selected for relating in a particularly clear way to one reason in question. In addition, an effort was made to diversify the selected studies with respect to where and when they were conducted.

3.1 Reason 1. Inform innovation design and development

This reason is germane to those studies that "examine questions about what the innovation could and/or should be, the extent to which an innovation is feasible in particular settings, and its utility from the perspective of the end users" (Century & Cassata, 2016, p. 174). It includes studies that "examine the creation of an innovation, its qualities and characteristics; understand place feasibility and usability of the innovation; or create an innovation customized for a time, and context" (p. 174).

A study by Confrey et al. (2017) can serve as a characteristic example. The study describes "the creation of a tool to meet the design challenge of scaffolding improved curricular coherence when practitioners use a variety of resources to build curriculum." (p. 732). This tool—a digital learning system called Math-Mapper (M-M) that includes visual maps of mathematical content called Relational Learning Clusters (RLC)—is *the innovation* in question. The paper first provides a detailed description of the innovation and its theoretical underpinnings. It then reports on three pilot studies where the innovation was enacted by the end-users—mathematics teachers and students from two school districts in the US—in order to test its feasibility and usability in practice. The first study explored how the M-M supported a process of curriculum revisions. A testimony of one of the teachers who "called her experience of the previous curriculum as 'chaotic' and celebrated the new one as 'calm'" (Confrey et al., 2017, p. 726) is presented as an illustration of feasibility and usability of the system. The second pilot study explored sixth grade students' performance on one of the RLCs, based on quantitative data collected from several hundred students. The third study explored the students' feedback after experiencing assessment that was organized in accordance with the M-M principles. Here the authors briefly summarized more than 1000 students' reflective reports and exemplified them through selected quotes. They cautiously concluded that the data "supports the validity of emphasis on learner-centeredness of curriculum coherence" (ibid., p. 732). Overall, in spite of the relatively broad scope of the research (two districts, tens of teachers, hundreds of students), the authors' final remark is that "the reported studies only offer an initial glimpse on M-M's potential effects on curriculum and instruction" (p. 732) and that further

studies are needed to explore outcomes of implementation, especially when conducted without ongoing support of the R&D team, and also to refine the design of the innovation.

Studies having a similar research focus—that is, studies that report on initial enactment of an innovation when its proponents are heavily involved—are not rare in the reviewed literature, though the nature of the innovations varies. Such studies concern, for example, innovative assessment systems (Ernest, 1984), electronic textbooks (Hoch et al., 2018), instructional guidelines (Colonnese, 2018), classroom pedagogies (Sullivan et al., 2013), instructional sequences (Stylianides & Stylianides, 2009), and professional development courses (Kuzle & Biehler, 2015).

3.2 Reason 2. If and how the innovation achieves desired outcomes for the target population

Studies in this group focus on examining the efficacy and effectiveness of innovations, and explore their emerging outcomes (Century & Cassata, 2016). This kind of research tries to determine whether the implementation of an innovation produces the expected results in the target population (students, teachers, schools, etc.)

Research of this kind was reported by Prediger et al. (2019). The study explored effects of combining three implementation strategies for upscaling professional development, namely, the community-based strategy, the material-based strategy and the systemic strategy. This combination, used in a professional development program, *Mastering Math,* in Germany, was the innovation at stake. The authors not only provided an existence proof of the viability of combining these strategies, but explored the effects of the innovation enactment on the participating teachers and students. The first two research questions of this study were as follows: (i) What are the effects of the research-based implementation and professional development program on teachers' perception of materials and of their cooperation in professional learning communities? (ii) What are the effects of the research-based implementation and professional development program on students' learning gains compared to a control group of students supported by the teachers outside the program? (p. 364). The data for the first question were collected via a questionnaire filled out by 63 teachers in the *Mastering Math* program. The data for the second question were collected using pre- and post-tests offered to about 600 students of the intervention group and to about 400 students of the control group. The results were encouraging at both teacher and student levels. However, the authors acknowledge methodological limitations of the study, and conclude by saying that "a combination of strategies *can* be effective" (p. 361, italics added).

Identification of 'effects' produced by innovations is an important and recurring theme in implementation research

in mathematics education. Effects are reported with qualifiers (e.g. time—sometimes, often; quantity—some, most) and with different degrees of confidence, and are subjects of many qualifications, but we suggest that Reason-2 studies can be grouped into the following four categories:

- Studies that focus on identifying effects of innovative teacher professional development programs and interventions (e.g., Beswick & Jones, 2011; Ferrini-Mundy et al., 2007); the aforementioned study by Prediger et al. (2019) belongs to this type.
- Studies related to effects of curriculum reforms and innovative curricula. For instance, there are studies that analyze effects of curriculum reform on students' knowledge and achievements depending on the level of fidelity of the implementation (e.g., Balfanz et al., 2006). There are also studies that investigate whether the implementation of curriculum innovations changes teachers' instructional practice (e.g., Obara & Sloan, 2010).
- Studies that analyze effects of implementing particular teaching approaches. Examples are as follows: (i) a study by Polotskaia and Savard (2018), who investigated effects of the relational paradigm on Canadian students' competencies in solving additive word problems; (ii) a study by Adam (2004) that explored outcomes of the implementation of an ethnomathematics curriculum unit for Maldivian students. This study relied on data collected from seven teachers and about 200 students, and reported an increase in measures related to motivation and interest, awareness of mathematics in society, and the understanding of mathematical concepts.
- Studies that analyze outcomes of the implementation of technological innovations. These studies are less common in the literature reviewed. An example is the work of Hoch et al. (2018) on the design and implementation of an electronic mathematics textbook, where researchers tried to identify an overall effect of 'time on task' on students' task success during initial instruction of fractions.

Reason-2 studies are highly diverse methodologically, ranging from relatively large-scope controlled experiments (as in Prediger et al., 2019), to relatively small-scope (mainly) qualitative studies (as in Adam, 2004). However, they all have a comparative component. For example, the study described by Adam (2004) does not include a control group, but the students were asked to indicate if they would prefer to study mathematics the way they learned in an ethnographic unit as implemented, or in a regular way. Generally speaking, the scope of the studies in this category varies, but they seem to conform to the maxim formulated by Adam (2004): "[t]he practical intervention was moderated by what was practically possible" (p. 57).

3.3 Reason 3. Understand relationships between influential factors, enactment, and outcomes

Reason 3 embraces studies that explore relationships between factors that can influence the innovation enactment and sometimes also attend to relationships between the contextual factors and outcomes of the implementation. These are studies that explore how, why, when, where, and with whom an innovation is effective (Century & Cassata, 2016, pp. 174–175). Identifying and understanding these factors are fundamental to implementation research, which makes Reason 3 one of the most frequently observed in the literature. Also, it is often the case that studies considered within Reason 3 are also included in some of the other reasons. In fact, the four categories used to classify the studies under Reason 2 can be used to organize the studies included in Reason 3 as well.

The work by Manouchehri and Goodman (2000) is a particularly salient illustration of Reason 3. It addresses the implementation of curriculum reforms and curricular innovations. In this work, the researchers used a qualitative case study to investigate the process of evaluation and implementation of an NCTM Standards-based textbook by two mathematics teachers over a period of two years. Their findings suggest that teachers' mathematical knowledge is one of the factors that most influences the way teachers evaluate and implement the textbook. Another illustration can be found in the study by Wright (2014) that analyzed the effectiveness of a teaching model as an instructional tool for the topic of percentages, ratio and proportions. As part of the results of this study, the researcher reported that the successful implementation of the teaching model "is dependent on the teacher noticing and responding to the layers of understanding demonstrated by students and the careful selection of materials, problems and situations." (p. 101).

In spite of the indicated overlap between Reason-2 and Reason-3 studies, there is also a characteristic difference: Reason-3 studies attend to individual differences among the participants and to contextual specificities of the implementation settings, whereas Reason-2 studies tend to report general effects. However, both Reason-2 and Reason-3 studies present implemented innovations as stable entities (e.g., a textbook or a curriculum unit), and are less explicit about iterative processes of the innovation design and re-design, as in studies driven by Reason 4.

3.4 Reason 4. Improve innovation design, use, and support in practice settings

Century and Cassata (2016) considered in this category studies that focus on improving the innovation and its implementation in order to improve outcomes as intended. Some

of these studies identify or develop supports needed for improving the use of innovations in practice. Other studies may instead be more focused on design iterations in order to improve the innovation itself.

The Reason-4 studies can be exemplified by the work of Clark-Wilson and Hoyles (2019). This study is part of a long-term research project in the UK that initially focused on the design of curriculum units embedding digital technology for learning mathematics. Later the project entered a phase of upscaling and dissemination. The researchers acknowledged the barriers and obstacles that may emerge when new end-users of the innovation join the project. As an action plan for overcoming the obstacles, they proposed to design a web-based 'professional development toolkit' that can support mathematics teachers' who implement the units in their classrooms beyond the timeline of the funded project. They also proposed the research basis for designing such a toolkit, including design principles.

Reason 4 manifests itself also in those studies that adopt a design research methodology with a focus on curriculum innovations. A characteristic example is the work by Kwon et al. (2015), where the researchers designed and implemented an inquiry-based multivariable calculus course containing various opportunities for students to discuss and argue. One of the goals of the study was to derive design principles, that is, to discover the characteristics of the instructional design that would supports students' argumentation. To achieve this goal, the authors adopted a design research methodology with iterative cycles comprising design, implementation, and reflection stages.

3.5 Reason 5. Develop theory

According to Century and Cassata (2016), this reason is characteristic of studies that aim to enhance our understanding of educational change by devising theories and frameworks. This type of study is not common in the reviewed mathematics education literature.

Reason-5 studies can be illustrated by the work of Maaß and Doorman (2013), in which a theoretical model for a widespread implementation of inquiry-based learning (IBL) is proposed. The authors acknowledge that it is not easy to change day-to-day teaching on a large scale, and therefore it is necessary to consider in depth the question of how to promote a widespread uptake of IBL in day-to-day teaching. To this end, Maaß and Doorman (2013) introduced a model for the dissemination and implementation of IBL. The authors explained the complexity of the model, including its theoretical basis, its iterative approach to evaluation and refinement, and its intended contributions to theory and practice.

The work by Jankvist and Niss (2015) is another example of Reason-5 research. These authors offer a framework for designing and implementing an in-service teacher education program in which findings from mathematics education research are put into practice. In particular, this program had the aim of helping teachers identify students with genuine learning difficulties in mathematics, investigate the nature of these difficulties, and carry out research-based interventions to assist the students in overcoming them.

One more example, already mentioned in Sect. 1, by Cobb and Jackson (2021), presented elements of *theory of action* and supported the claim of its usability in the context of a large-scale program in the US. Additional examples of this type are overviewed as part of this SI.

3.6 Intermediate remarks

In conclusion of this section, we would like to make three remarks. ***First***, it should not be surprising that mathematics education implementation research, identified as such in accordance with Century and Cassata's (2016) working definition, embraces studies that could be, and actually are, characterized by their authors as design research, intervention studies, controlled experiments, teaching experiments, etc. However, not every design research or controlled experiment would qualify. It is the authors' decision to study enactment of a resource/innovation/approach not only in the context in which it was originally developed but in a new context, that makes the study fit Century and Cassata's definition. ***Second***, some authors self-identify their studies as 'background research' or as 'research accompanying an implementation project', but rarely can we find 'implementation research' as an explicitly stated type of research. This is in contrast to some fields of study other than mathematics education (cf., Century & Cassata, 2016; Eccles & Mittman, 2006). In light of this observation, we choose to use the term 'implementation-related research' rather than 'implementation research' henceforth. The ***third*** remark is that studies on long-term consequences of implementation of innovations—that is, what happens when the innovation proposers step out and the innovation remains essentially in the hands of its end-users—are extremely rare in the reviewed mathematics education literature.

4 Implementation-related research in this special issue

4.1 A general overview of the SI

Classifying the SI papers according to Century and Cassata's (2016) five reasons for conducting implementation research turned out to be not especially informative. This is because most of these papers simultaneously encompass several reasons. For example, Jaworski and Potari (2021) mentioned all five reasons as the motivation for their study,

and many authors developed arguments that concern four reasons (e.g., Burkhardt & Schoenfeld, 2021; Krainer, 2021; Prytz, 2021) or three (e.g., Devlin, 2021; Jankvist et al., 2021a, 2021b; Valoyes-Chávez & Felmer, 2021; Wang et al., 2021). In other words, most of the papers simultaneously elaborate on design and/or re-design issues, on the interplay of factors that influence the enactment and on aspects of implementation-related theory development. Of course, the multitude of the reasons for conducting research does not mean that the SI papers are unfocused. It is rather indicative of the specific setting of this SI, which was conceived as a platform for the collective search for identity of implementation as a sub-domain in mathematics education research (cf. a search for the identity of mathematics education as a research domain described in the book edited by Sierpinska & Kilpatrick, 1998). It can also be an indication of the mathematics education research tradition that treats design, enactment and transfer across contexts as interwoven and theory-laden (Adler et al., 2005; Niss, 1996, 1999).

As could be expected in a rapidly developing but not yet institutionalized field of study, the SI papers contain diverse suggestions about what 'implementation' and 'implementation research' are or should be in mathematics education. Some authors situate their work in one of the existing conceptualizations of 'implementation' and 'implementation research' (see Sect. 1) whereas others challenge some of the basic assumptions on which the existing conceptualizations rely. The unifying feature is that all the papers describe and explore efforts made in pursuit of some educational change in natural (or fairly natural) mathematics education habitats. In this way, the SI papers come close to ecological perspectives, which consider innovations as "ecological disruptions for the didactic system in which they are implemented" (Artigue, 2021, p. 31).

With respect to the types of argument the authors construct in order to make their main points, six paper can be tagged as *theoretical*, i.e., papers in which a theoretical argument representing the authors' position is construed and then supported by illustrative examples taken from the authors' past studies, and nine papers that can be tagged as *empirical*, in which the original data sets are analyzed within a particular implementation-related theoretical perspective.

As explained in Sect. 2, the reading of the past literature and the SI papers led us to identification of four themes that we deem particularly important for further discussion in our community. Each theme represents diversity and even controversy in the authors' approaches, which allows us to seize a valuable opportunity to map the implementation-related mathematics education research landscape (as represented by this particular collection of papers), by comparing and contrasting different approaches. These themes are developed in Sects. 4.2–4.6.

4.2 Objects of implementation

For Century and Cassata (2016), the object of implementation is *innovation*, which is broadly defined as "programs, interventions, technologies, processes, approaches, methods, strategies, or policies that involve a change (e.g., in behavior or practice) for the individuals (end users) enacting them" (p. 170). This notion is succinctly denoted as "the it" or "the focus of change" (ibid). In the context of this SI, characteristics of objects of implementation are important as they seem to relate strongly to preferences in research. Two characterizations are in order. The first is related to the medium of an innovation, that is, the form in which an innovation 'exists' or is proposed so that it could be implemented by someone. The second is related to the extent to which 'the focus of change' is predefined (by the proponents) or co-constructed (by the proposers and the adapters).

As for the first characterization, we see (i) innovations, in which know-how is encapsulated in a relatively stable material artefact around which the human activity is developed (e.g., a collection of published instructional materials), and (ii) innovations which mainly exist in the form of human interactions driven by ideas or principles (e.g., a community of inquiry).[2] Innovations of the second type can be supported by material artefacts but do not fully depend on them; they rather depend on a communicational infrastructure and rules of interaction between the involved individuals and communities. We refer to these two types of objects of implementation as *material-centered* and *interaction-centered*, respectively.

Examples of *material-centered objects of implementation* are present in the papers by Burkhardt and Schoenfeld (e.g., a collection of instructional materials on the website of the Mathematics Assessment Program) and Karsenty (a collection of recorded mathematics lessons on the website of the VIDEO-LM project). *Interaction-centered objects of implementation* are described in the papers by Jaworski and Potari (the developmental model for enhancing an inquiry stance), Pinto and Koichu (practices and processes of disciplined inquiry), Roesken-Winter et al. (insights from past research with teachers and facilitators conducted by the authors), among others.

Of course, there are also implementation programs, in which nearly symmetrical attention is given to both material-centered and interaction-center innovations that complement each other (e.g., Diego-Mantecón et al., 2021; Jankvist et al.,

---

[2] Admittedly, this categorization resonates with Adler's (2000) categorization of resources in school mathematics. Adler distinguishes between human resources (such as teacher–pupil ratios, class size, teacher qualifications), material resources (such as technologies, textbooks, mathematical objects and everyday objects) and social and cultural resources (such as language and time).

2021a, 2021b; Karsenty, 2021; Roesken-Winter et al., 2021; Wang et al., 2021). Such programs concern innovations that can be seen as mixed or *material-interactive*. Existence of such programs implies that the above dichotomy is blurred. However, we deem it useful because it can partially account for certain differences in the authors' foci of research attention. Namely, studies on implementation of material-centered innovations more readily discuss implementation outcomes, whereas studies focusing on interaction-centered innovations are more deliberate about processes of implementation. In addition, it seems that material objects of implementation favor the 'improvement of practice at scale' problematique, while interactional objects of implementation favor 'bridging research and practice' agendas.

And what about mathematics education research as an object of implementation? Overall, this seems to belong to the interaction-centered type of object, or to the mixed material-interactive type, but never to the material-centered type only. Actually, research as an object of implementation is alluded to in many studies as a conjunction of approaches, resources, organizational models or theories developed in the past within an organizational frame called 'research' (e.g., Burkhardt & Schoenfeld, 2021; Wang et al., 2021). Some authors avoid the use of the collocation 'implementation of research' and talk about 'implementation of research-based innovations' (e.g., Jaworski & Potari, 2021; Roesken-Winter et al., 2021). The other authors, however, adhere to the 'implementation of research' language. For example, Pinto and Koichu (2021) talk about implementation of research as teachers' engagement with procedures and constructs of disciplined educational inquiry. The engagement occurs in the context of a community of inquiry and is mediated by a set of boundary objects. Some of these objects are material (e.g., a written guide for data-collection and analysis), and others are purely interactive (e.g., practices developed in the community for making decisions about further action). Kontorovich and Bartlett (2021) and Cai and Hwang (2021) describe implementation of research as incorporating in teaching tasks of particular types that have been extensively explored in the past, namely, scriptwriting tasks and problem-posing tasks, respectively. In these studies, teacher-researcher interactions towards developing a predisposition for the use of the chosen types of tasks are analyzed in depth, and material artefacts supporting implementation (e.g., specific tasks) are co-constructed in due course of the interactions.

The second characterization of objects of implementation—by whether an innovation as 'the object of change' is predefined and stable or gradually co-constructed and flexible—results in quite a different classification. For example, the TBM project presented by Jaworski and Potari (2021) had a clear predefined goal, namely, to promote inquiry as a way of being, by means of implementing at scale the developmental PD model developed in prior research. Likewise, a project described by Tamborg (2021) aimed to achieve a pre-defined goal, namely, to promote an objective-oriented approach to teaching (i.e., an approach in which teaching is driven by well-defined learning objectives) by means of making a particular digital platform mandatory for all teachers of the country. Both projects have pre-defined goals but differ in the types of their objects of implementation, which were interaction-centered and material-interactive respectively. Interestingly, both studies focus on tensions and pitfalls of implementation. In contrast, the studies, in which 'the object of change' is not pre-defined but co-constructed under a flexible theoretical umbrella (e.g., Cai & Hwang, 2021; Valoyes-Chávez & Felmer, 2021) are essentially described as 'success stories', in the following meaning: they extensively report on learning gains in the process of implementation (and sometimes also in its outcomes; see Wang et al., 2021), and focus less on tensions and pitfalls of the implementation processes. This leads us to a paramount question, addressed in the next section: who decides on what 'the object of change' should or can be, and on what is this decision based?

4.3 Configurations of stakeholders in implementation

Krainer (2021) presents the following four idealized implementation scenarios representing four different configurations of stakeholders: (i) teachers identify a practical problem and search for the solution—policy makers and researchers are not involved, local implementation is successful but scaling up is hardly possible; (ii) researchers propose an innovation—a group of teachers successfully implement it, policy makers are informed of the success but scaling up is uncertain; (iii) policy makers identify a systemic problem—researchers advise, there are disagreements about focus, timescale and resources, action is uncertain; (iv) a policy maker and a practitioner consider an innovative program while trying to address different problems—there are issues related to scope, expectations, timescale and operation, action is initially impossible, more thinking, possibly including researchers, is needed.

Three types of stakeholders—practitioners, researchers, policy makers—interact in these scenarios, as they do in the papers of this SI. Nevertheless, some scenarios are more commonly addressed in research than others. The first scenario, describing a situation where practitioners initiate educational change and pursue it by themselves, is absent in this SI and in the reviewed literature. This is not to say that such situations cannot exist in practice: perhaps they exist but are not documented. The same holds for scenario (iv). However, two studies of the SI are quite close to scenario (i). Cai and Hwang (2021) describe a teacher who is

interested in improving her lessons and seeks collaboration with a researcher, who proposes that the teacher adapt a research-based innovation. In the second case, Kontorovich and Bartlett (2021) describe a researcher who approaches a practitioner with an idea for an innovation while pursuing his research agenda, and collaboration becomes possible only when the researcher adjusts his research agenda to the agenda of the practitioner, who eventually implements the innovation for the sake of addressing a teaching need. In both cases, scaling up of the implementation is not considered as realistic. Scenario (ii) is present in the study by Pinto and Koichu (2021): researchers offer an innovative PD program, in which experienced teacher-participants enter the shoes of education researchers and implement the innovation—teaching aligned with the educational research cycle—in order to improve aspects of their teaching. Scaling up is uncertain. Implementation initiatives described by Burkhardt and Schoenfeld (2021), Devlin (2021), Diego-Mantecón et al. (2021), Jaworski and Potari (2021), Karsenty (2021) and Roesken-Winter et al. (2021) partially conform to scenario (ii), in that the change is initiated by researchers but not necessarily in collaboration with policy makers, who enter the picture later on. Scaling up is achieved to different extents. Finally, scenario (iii), in which policy makers initiate the change, on their own or in collaboration with researchers, appears in papers by Krainer (2021), Tamborg (2021), Valoyes-Chávez and Felmer (2021) and Wang et al., (2021). Not surprisingly, projects in which policy makers are involved from the beginning always aim to implement at scale, projects in which researchers are proponents of innovations vary in this respect, and projects initiated by teachers are neutral with regard to upscaling.

Furthermore, in three SI papers, by Jankvist et al., (2021a, 2021b), Krainer (2021), and Prytz (2021), the complexity of relationships between the stakeholders is *the* topic of investigation. Here we would like to briefly dwell on the study by Prytz (2021), which reveals not only the complexity, but also dynamics of the various roles of stakeholders in the long run. While exploring the history of mathematics education reforms in Sweden, Prytz describes how the role of researches has evolved from proponents of innovations to be adapted by teachers in a highly centralized educational system in the 1960s, to supporters and explorers of teacher-initiated innovations in a highly decentralized educational system in the 2010s.

Last but not least, innovations are conceived and enacted not out of the blue but in order to resolve what is seen by particular stakeholders as an acute problem (Bryk et al., 2015; Krainer, 2021). We tentatively observe that implementation projects initiated by researchers, practitioners or policy makers tend to have different justifications. Researcher-initiated projects tend to put forward theoretical justifications of proposed innovations (e.g., we know that

Problem Based Learning in transdisciplinary contexts is important, so let's implement it, as in the work of Diego-Mantecón et al., 2021). Practitioner-initiated projects stem from the description of a problem of practice (e.g., a teacher is unsatisfied with her students' achievements and motivation, as in Cai & Hwang, 2021). Policy-maker initiatives usually stem from a systemic problem (e.g., poor alignment of teaching with the objectives identified as important by officials, as in Tamborg, 2021). This SI contains only one example where all three stakeholders are involved and seem to co-exist in harmony: it is the study of the *Just Do Math* program in Taiwan (Wang et al., 2021), which during several years reached the national level and was considered successful by various measures. In this case, communication channels among different stakeholders were carefully designed as part of the overarching research-driven model of implementation from the very beginning.

4.4 Implementation vs. scaling up

Implementation is often interpreted as the application in practice of an innovation generated by a group of specialists, with the aspiration of disseminating the innovation and scaling it up (Maaß & Doorman, 2013). A conceptualization of implementation reflecting these perspectives was provided by Maaß and Artigue (2013): "Implementation is what happens when a planned intervention, an innovation, is set in motion. When an intervention is designed, its designers often do not aim merely for a small-scale implementation, but also wish to disseminate their ideas, materials, etc." (p. 779).

Judging from scenarios and examples discussed in Sect. 4.3, implementation and scaling up can be seen as related but not synonymous notions. They both presume that an innovation designed by a group of specialists in a particular context is then enacted in new contexts, with or without direct involvement of the original specialists. The difference is subtle: scaling up presumes that the new contexts are larger (in terms of numbers of students, teachers, schools or districts) than the initial context, whereas implementation is concerned with how an innovation works *across contexts*, as Cai and Hwang (2021) suggest.[3] The former approach always values dissemination, and the latter approach *can* value dissemination, but also some other important goals, such as resolving teaching problems in particular classrooms. Let us refer to the first approach as *implementation*

---

[3] A related discussion can be found in the editorial by Cai et al. (2020), in which they introduce the *replimentation* notion. Replimentation refers to the duality of goals that researchers may pursue: to implement an innovation in new contexts to scale it up, and to treat the new implementations as replications with variations to explore the innovation.

at scale, and to the second one as *local implementation* (this is to follow 'classroom implementation' introduced by Cai et al., 2017).

Implementation at scale is at the center of most of the papers in this SI. However, what counts as a worthy scale of implementation is very idiosyncratic. For example, Jaworski and Potari (2021) describe an innovation (the developmental model) conceived and designed in the context of a learning community comprising 12–15 didacticians and about 30 teachers, which was then enacted in five districts in Norway. Devlin (2021) describe gamified mathematical activities designed by a small R&D team that are eventually used by hundreds of thousands of students over the world. Roesken-Winter et al. (2021) describe an 18-year-long *Mastering Math* project, in which the innovation (research-based curriculum materials) was developed in-house over six years, was enacted in first three schools, then in 40 schools, and eventually in more than 200 schools. Prytz (2021) describes a developmental project involving 400 teachers followed by the large-scale *Boost* project involving the majority of mathematics teachers in Sweden. He concludes: "[w]hat is possible with 400 teachers in one municipality is perhaps not possible with 76 percent of all Swedish mathematics teachers" (this issue). Karsenty (2021) reports dissemination of the project *SHLAV* as follows: it began from 2 schools (6 teachers, 100 students) in the pilot phase and extended to 32 schools (191 teachers and 4759 students) in the implementation phase. Burkhardt and Schoenfeld (2021) describe the impact of the *Connected Mathematics* project in terms of its dissemination among practitioners and researchers: apparently (the exact numbers are not available), they talk about thousands of teacher-users and hundreds of researcher-users.

Not surprisingly—a similar observation was made by Venkat and Graven (2017) regarding material vs. human PD resources—material-centered objects of implementation can be scaled up more readily than interaction-centered objects of implementation. However, what is striking in the diversity of scope attained in the abovementioned projects is that different scholars seem to assign very different meaning to implementation as 'enactment' or 'setting in motion'. We have already observed that types of implementation differ depending on the type of the implementation objects (Sect. 4.2). To this we now add an observation informed by the distinction made by Karsenty (2021): there are two kinds of scaling up—of the setting when the program is implemented in new contexts as a package to be used essentially 'as is', and of values, where practitioners, stimulated by the program's ideas, attain full agency over how to use these ideas in practice. Thus, 'enactment' as a phenomenon is tied to the degree of professional autonomy of its users, as well as to the degree to which proponents of an innovation feel that credit belongs to them even when the innovation is essentially modified or used in unexpected ways, or

is used in conjunction with other innovations developed in other projects.[4] Further theorizing of such notions as 'enactment' and 'setting in motion' is needed as part of theorizing 'implementation at scale' in mathematics education.

We now turn to examples of local implementation in this SI. Four studies, namely those by Cai and Hwang (2021), Diego-Mantecón et al. (2021), Kontorovich and Bartlett (2021) and Pinto and Koichu (2021), belong to this category. The unifying feature of these studies is that they operate with *interaction-based objects of implementation* (see Sect. 4.2), which are deeply rooted in specific insights from focused sub-domains of mathematics education research. (To recall, the sub-domains are problem posing in Cai and Hwang's case, Problem Based Learning in transdisciplinary contexts in Diego-Mantecón et al.'s case, scriptwriting tasks in Kontorovich and Bartlett's case, and teacher inquiry with a special focus on empowering school students to ask good questions in mathematics lessons in Pinto and Koichu's case.)

A natural question is, what added value can there be in considering these small-scale studies as instances of implementation-related research? Moreover, would it not be more natural to consider them as design studies or teaching experiments, given that the methodologies applied in these studies bear clear signs of these types of research? Our response is two-fold. First, we would like to reiterate (see Sect. 3) that it is not a particular methodology, but rather specific research questions focusing on 'implementation' that matter. Second, we would like to rely on Burkhardt and Schoenfeld's (2021) assertion: "In education, it is rare for a single research result to form the basis for a change in practice" (this issue). We agree that it is rare, and value that this is the case in the four studies in question. We deem that dealing with the challenge of changing practice based on specific research results is what gives merit to these studies within the implementation problematique. Speaking more generally, some changes in practice are so desirable and simultaneously so difficult to achieve, that even a small-scope implementation is of academic value, especially if it is documented in detail and can inform the field in ways germane to in-depth qualitative research. For example, the study by Kontorovich and Bartlett (2021) contains a fine-grained analysis of interactions between the mathematics education researcher and the mathematics lecturer that result in the inclusion of an unorthodox assignment in a lecture-based mathematics course for undergraduate students. This is a rare achievement. The analysis cannot be generalized directly, but it is theorized in the paper and therefore can be useful for constructing

---

[4] Admittedly, the second part of this observation (about the innovation proposers) is inspired by Rogers' (2003) discussion of phenomena pertinent to diffusion of innovations and by Remillard's (2005) observation that 'fidelity' is a problematic concept when applied to curriculum use.

researcher-practitioner interactions in future projects, including large-scale ones.

Speaking on the benefits of small-scale qualitative studies, it is of note that most of the large-scale implementation projects mentioned above are accompanied by this type of research. For example, the *Mastering Math* project (Roesken-Winter et al., 2021) in Germany was informed and supported by a series of small-scale design and intervention studies. A national-level implementation project in Denmark presented by Tamborg (2021) was analyzed in light of the data collected from a group of seven teachers. An aspect of a large-scale *APRA* project in Chile was explored by Valoyes-Chávez and Felmer (2021) by means of a case study of one participant. Furthermore, the data corpus on the *Teaching Better Mathematics* project in Norway that Jaworski and Potari (2021) present consists of reflections of the leading didacticians of the project. Jankvist et al., (2021a, 2021b) and Prytz (2021) explored the long-term reform of mathematics education in Sweden as an educational case study that relies on the analysis of relevant documents and insights from past studies. As a matter of fact, in this SI only the national-level *To-Do-Math* project in Taiwan (Wang et al., 2021) is accompanied by systematic analysis of quantitative data collected at a scale comparable with the scale of the project.

The following trend seems to emerge: there is often a gap between the scope of the implementation projects and the scope of the related research. The gap is fully understandable: as a rule, mathematics education research is conducted by relatively small research teams having limited capacity for collecting and processing large-scale data, and especially large-scale qualitative data. Paraphrasing Adam's (2004) maxim quoted in Sect. 3.2, even in large-scale projects research is restricted by what is practically possible. Simultaneously, we recognize that establishing normative relationships between the project-scope and the research-scope requires more thinking. On the one hand, not everything in large-scale projects should and can be documented and explored. On the other hand, the role of the large scope of a project in small-scale studies accompanying the project should remain visible.

4.5 Implementability of mathematics education research

Research-based innovations differ in their *implementability*. Drawing on Levine and Cooper (1991), Jankvist et al., (2021a, 2021b) refer to *implementability* of a research-based innovation in an educational context as an indicator of "how realistic and feasible it is for practitioners to implement an innovation arising from *interpreted* research results" (this issue, italics added). In turn, interpretation of research results is necessary in order to make them accessible for the

intended end users. Therefore, implementability of research-based innovations depends on who interprets and mediates them, and how.

We have already discussed mathematics education research as an object of implementation in Sect. 4.2. We now unfold that discussion by considering implementability of mathematics education research products, including research findings, theoretical frameworks and experiences (Aguilar et al., 2019).

As for specific *research findings*, we have discussed that their implementation is possible but rare (Sect. 4.4). In three cases of local implementation included in this SI (Cai & Hwang, 2021; Kontorovich & Bartlett, 2021; Pinto & Koichu, 2021), special attention was given to artefacts supporting implementation. Cai and Hwang (2021) give special merit to tangible artefacts that emerge from teaching. These artefacts, they argue, can mediate between research and practice. In the two other studies, the *boundary object* notion (e.g., Star, 2010) is put forward. Studies in the implementation-at-scale category, which usually do not operate with specific findings but with clusters of findings developed in realms of mathematics education research (e.g., research on conceptual learning), are also attuned to mediation of research findings through artefacts. For example, Roesken-Winter et al. (2021) discuss how material implementation strategy (i.e., the use of research-based instructional materials with students, teachers and facilitators) can support the scaling up processes. In sum, implementability of research findings seems to be closely related to ways by which these findings are interpreted and transposed into artefacts accessible to end users. Last but not least, mediation of research findings in the form of advice to policy makers is also considered, for example, by Krainer (2021), but implementability of this sort of mediation is uncertain.

Unfolding implementability of *theoretical frameworks* seems to be even trickier than the implementability of research findings. Though all projects and studies in this SI are driven by theories and theoretical frameworks, no claims are made for their (direct) implementation. While not surprising, it is a phenomenon that requires interpretation. Ours stems from a seemingly paradoxical idea of Mason about mathematics education research as an enterprise. Mason (1998, 2002) suggested that the main enterprise of mathematics education researchers is to learn about themselves through interactions with others (cf. Sect. 1 on the improvement mathematics education as the *raison d'être* for the existence of mathematics education research). To extend Mason's idea, one can argue that theories and conceptual frameworks deeply affect their creators and sometimes additional researchers who are willing to study them and are trained to put them into action. However, the influence of mathematics education theory on other stakeholders of implementation (i.e., practitioners and policy makers) is

hindered for many reasons (see, e.g., Jankvist et al., 2021a, 2021b). Practically, in light of the current view of implementation as a complex endeavor based on communication of many parties, and also in light of the call to make implementation an integral part of research (Cai & Hwang, 2021; Cai et al., 2017), creators of theories might consider how to coin their concepts and constructs so that at least some of these would be communicable beyond the research community. Otherwise, the evident complexity of conceptual apparatuses used by mathematics education researchers may pose yet another obstacle for implementability of research.

Finally, implementability of *research experiences* is considered in two studies of the SI. It is the main topic of the study by Pinto and Koichu (2021), and an aspect of the developmental model of Jaworski and Potari (2021). In particular, Pinto and Koichu's (2021) study shows how the border between practices of disciplined inquiry (i.e., research practices) and of teacher inquiry can be blurred when both parties productively collaborate over specially designed boundary objects without an expectation of reaching full consensus. Implementability of research experiences, is, however, a peripheral topic in this SI, though it is extensively discussed elsewhere (e.g., Cobb, 2000; Kieran et al., 2012).

5 Concluding remarks

We have engaged with a growing body of mathematics education studies that relates to the implementation of innovations in a more or less direct manner. To make sense of this body of studies, we followed a hermeneutical approach in which five reasons for developing implementation research (Century & Cassata, 2016) were instrumental for the initial organization of the literature. The hermeneutic approach, applied to past research and to the studies included in this SI, allowed our understanding to evolve towards the identification of four themes that we deem important for further research and practice, namely, objects of implementation (two categorizations are offered), stakeholders in implementation (possible configurations are illustrated), emerging differences between implementation and scaling up (a distinction between dissemination and implementation across contexts is substantiated), and implementability of mathematics education research (implementability of different research products is discussed).

In what follows, we propose a refined conceptualization of the notion of *implementation in mathematics education*, which may serve as an organizational framework for future research, by signaling the key actors and processes within this enterprise. In addition, a conceptualization of implementation-related research is also offered in conjunction with suggestions for future research.

We conceptualize *implementation in mathematics education* as an ecological disruption to a particular mathematics education system, through the gradual endorsement of innovation in conjunction with an action plan aimed at resolving what is perceived as a problem by (at least some of) the stakeholders involved. The defining feature of implementation is that it occurs in interaction between *the innovation and plan proponents* and *the innovation adapters*. At the beginning of the implementation, the innovation proponents have the ultimate agency over the innovation and the associated action plan. During the implementation process, the innovation adapters experience some or all of the following sub-processes: (1) constructing agency over the innovation, (2) gradually changing within-community communication or across-community communication, (3) gradually changing practice so that it accommodates the innovation, (4) adapting the innovation to their needs and aspirations. These sub-processes reflect back on the proponents, including evolution of the innovation, of the associated action plan and of the theories underlying their development. The implementation process is iterative and ends when the innovation stops being perceived by the stakeholders as an ecological disruption. To this end, implementation can succeed (e.g., the adapted innovation is eventually integrated in the system) or fail (i.e., the innovation is rejected explicitly or tacitly so that it stops influencing the system).

While recognizing the specificity of mathematics education as a field of study, we prefer to talk about *implementation-related research* (IRR) rather than on implementation research. We refer to IRR as a disciplined inquiry of implementation (as specified above) that aims at creating new theoretical and practical knowledge on the use of innovations beyond contexts in which they have been created.

Given the apparent inclusiveness of the above conceptualization and the broadly recognized fact that 'good research' presumes that its results are to be used by someone or for something, one may wonder what is *not* implementation and what is *not* IRR. Well, quite a few types of examples can be constructed and illustrated, by negation of one or more of the above-listed characteristic features of implementation. However, it might be more productive to ask what the core and future direction of IRR may be. The papers in this SI suggest a focus on understanding and creating alignment between various stakeholders' different goals in relation to implementation. Researchers in IRR should be aware of what types of goals are in play in relation to the implementation problem (Niss, 1996), by whom these goals are articulated and what rationality they are based on (Krainer, 2021). For now, we see developing epistemic tools for better understanding the alignment between implementation stakeholders as one of the key concerns for IRR as it moves forward.

In addition, we see the pursuit of another type of alignment—between theoretical resources used in mathematics

education IRR studies and theoretical resources developed in other fields of study or in other sub-domains of mathematics education research (e.g., curriculum use)—as another important developmental direction for IRR. This type of alignment work has been visible in CERME10 and CERME11 (e.g., Jankvist et al., 2019), is continued in this SI, and, hopefully, will be continued on pages of the aforementioned newborn journal IRME, *Implementation and Replication Studies in Mathematics Education (*Artigue, 2021; Cobb & Jackson, 2021; Jankvist et al., 2021a), among other platforms.

Two recent initiatives, this SI and IRME, arose from the discussions and collective reflections that took place at CERME11. As mentioned, the goal of this SI was to foreground further the implementation problem in our community. The goal of IRME is to provide a stable space for academic interaction that could further support research-based improvement of mathematics education. Since we, the authors of this paper, are involved in both initiatives, we would like to finish this paper with an open invitation to mathematics education practitioners, researchers, didacticians and policymakers to use IRME as a platform for continuing the discussion of the implementation problem in mathematics education.

References

Adam, S. (2004). Ethnomathematical ideas in the curriculum. *Educational Studies in Mathematics, 16*(2), 49–68. https://doi.org/10.1007/BF03217395

Adler, J. (2000). Conceptualising resources as a theme for teacher education. *Journal of Mathematics Teacher Education, 3*(3), 205–224. https://doi.org/10.1023/A:1009903206236

Adler, J., Ball, D., Krainer, K., Lin, F.-L., & Novotna, J. (2005). Reflections on an emerging field: Researching mathematics teacher education. *Educational Studies in Mathematics, 60*(3), 359–381. https://doi.org/10.1007/s10649-005-5072-6

Aguilar, M. S., Kuzle, A., Wæge, K., & Misfeldt, M. (2019). Introduction to the papers of TWG23: Implementation of research findings in mathematics education. In U. T. Jankvist, M. van der Heuvel-Panhuizen, & M. Veldhuis (Eds.), *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education* (pp. 4355–4362). Freudenthal Group & Freudenthal Institute; Utrecht University; ERME.

Artigue, M. (2021). Implementation studies in mathematics education: What theoretical resources? *Implementation and Replication Studies in Mathematics Education, 1*(1), 21–52. https://doi.org/10.1163/26670127-01010002

Balfanz, R., Mac Iver, D. J., & Byrnes, V. (2006). The implementation and impact of evidence-based mathematics reforms in high-poverty middle schools: A multi-site, multi-year study. *Journal for Research in Mathematics Education, 37*(1), 33–64. https://doi.org/10.2307/30035051

Beswick, K., & Jones, T. (2011). Taking professional learning to isolated schools: Perceptions of providers and principals, and lessons for effective professional learning. *Mathematics Education Research Journal, 23*(2), 83–105. https://doi.org/10.1007/s13394-011-0006-3

Blomhøj, M. (2021). Samspil mellem fagdidaktisk forskning og udvikling af matematikundervisning—belyst gennem erfaringer fra et udviklingsprojekt i undersøgende matematikundervisning [Interaction between didactic research and development of mathematics teaching—Viewed through experiences in a development project in exploratory mathematics teaching]. *Sammenlignende Fagdidaktik, 6*, 29–50. https://doi.org/10.7146/sammenlignendefagdidaktik.vi6.127569

Boell, S. K., & Cecez-Kecmanovic, D. (2014). A hermeneutic approach for conducting literature reviews and literature searches. *Communications of the Association for Information Systems, 34*, 257–286. https://doi.org/10.17705/1CAIS.03412 Article number 12.

Bruckheimer, M. (1979). Creative implementation. In P. Tamir, A. Blum, A. Hofstein, & N. Sabar (Eds.), *Proceedings of the Bat-Sheva Seminar on curriculum implementation and its relationships to curriculum development in science* (pp. 43–49). The Weizmann Institute of Science.

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to improve: How America's schools can get better at getting better*. Harvard Education Press.

Burkhardt, H., & Schoenfeld, A. (2021). Not just "implementation": The synergy of research and practice in an engineering research approach to educational design and development. *ZDM - Mathematics Education*. https://doi.org/10.1007/s11858-020-01208-z

Burkhardt, H., & Schoenfeld, A. H. (2003). Improving educational research: Toward a more useful, more influential, and better-funded enterprise. *Educational Researcher, 32*(9), 3–14. https://doi.org/10.3102/0013189X032009003

Cai, J., & Hwang, S. (2021). What does it mean to make implementation integral to research? *ZDM - Mathematics Education, 53*(5). https://doi.org/10.1007/s11858-021-01301-x

Cai, J., Morris, A., Hohensee, C., Hwang, S., Robison, V., Cirillo, M., Kramer, S., & Hiebert, J. (2020). Working across contexts: Scaling up or replicating with variations? *Journal for Research in Mathematics Education, 51*(3), 258–267. https://doi.org/10.5951/jresemtheduc-2020-0007

Cai, J., Morris, A., Hohensee, C., Hwang, S., Robison, V., & Hiebert, J. (2017). Making classroom implementation an integral part of research. *Journal for Research in Mathematics Education, 48*(4), 342–347. https://doi.org/10.5951/jresematheduc.48.4.0342

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

Choppin, J., McDuffie, A. R., Drake, C., & Davis, J. (2018). Curriculum ergonomics: Conceptualizing the interactions between

curriculum design and use. *International Journal of Educational Research, 92*, 75–85. https://doi.org/10.1016/j.ijer.2018.09.015

Clark-Wilson, A., & Hoyles, C. (2019). A research-informed web-based professional development toolkit to support technology-enhanced mathematics teaching at scale. *Educational Studies in Mathematics, 102*(3), 343–359. https://doi.org/10.1007/s10649-018-9836-1

Cobb, P. (2000). Conducting teaching experiments in collaboration with teachers. In A. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 307–333). Lawrence Erlbaum.

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

Colonnese, M. W. (2018). The development of instructional guidelines for elementary mathematical writing. *School Science and Mathematics, 120*(3), 129–143. https://doi.org/10.1111/ssm.12391

Confrey, J., Castro-Filho, J., & Wilhelm, J. (2000). Implementation research as a means to link systemic reform and applied psychology in mathematics education. *Educational Psychologist, 35*(3), 179–191. https://doi.org/10.1207/S15326985EP3503_4

Confrey, J., Gianopulos, G., McGowan, W., Shah, M., & Belcher, M. (2017). Scaffolding learner-centered curricular coherence using learning maps and diagnostic assessments designed around mathematics learning trajectories. *ZDM - Mathematics Education, 49*(5), 717–734. https://doi.org/10.1007/s11858-017-0869-1

Devlin, K. (2021). Creating an educational tool that uses interactive representations to provide a new approach to mathematics learning. *ZDM - Mathematics Education*. https://doi.org/10.1007/s11858-021-01235-4

Diego-Mantecón, J.-M., Prodromou, T., Lavicza, Z., Blanco, T., & Oritz-Laso, Z. (2021). Are STEAM project-based learning implementations suitable to mathematically prepare our teenagers for the current society? *ZDM - Mathematics Education*, 53(5).

Eccles, M. P., & Mittman, B. S. (2006). Welcome to *Implementation Science*. *Implementation Science*. https://doi.org/10.1186/1748-5908-1-1

Ernest, P. (1984). The CXC mathematics examination: A Caribbean innovation in assessment. *Educational Studies in Mathematics, 15*(4), 397–412. https://doi.org/10.1007/BF00311114

Ferrini-Mundy, J., Burrill, G., & Schmidt, W. H. (2007). Building teacher capacity for implementing curricular coherence: Mathematics teacher professional development tasks. *Journal of Mathematics Teacher Education, 10*(4–6), 311–324. https://doi.org/10.1007/s10857-007-9053-9

Hoch, S., Reinhold, F., Werner, B., Richter-Gebert, J., & Reiss, K. (2018). Design and research potential of interactive textbooks: The case of fractions. *ZDM - Mathematics Education, 50*(5), 839–848. https://doi.org/10.1007/s11858-018-0971-z

Jankvist, U. T., Aguilar, M. S., Ärlebäck, J. B., & Wæge, K. (2017). Introduction to the papers of TWG23: Implementation of research findings in mathematics education. In T. Dooley & G. Gueudet (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education* (pp. 3769–3775). DCU Institute of Education; ERME.

Jankvist, U., Aguilar, M. S., Dreyøe, J., & Misfeldt, M. (2019). Adapting implementation research frameworks for mathematics education. In U. Jankvist, M. van den Heuvel-Panhuizen, & M. Veldhuis (Eds), Proceedings of the *Eleventh Congress of the European Society for Research in Mathematics Education* (pp. 4405–4412). Freudenthal Group & Freudenthal Institute; ERME.

Jankvist, U. F., Aguilar, M. S., Misfeldt, M., & Koichu, B. (2021a). Launching Implementation and Replication Studies in Mathematics Education (IRME). *Implementation and Replication Studies*

in Mathematics Education, 1(1), 1–19. https://doi.org/10.1163/26670127-01010001

Jankvist, U. T., Gregersen, R. M., & Lauridsen, S. D. (2021b). Illustrating the need for a 'Theory of Change' in implementation processes. *ZDM - Mathematics Education*. https://doi.org/10.1007/s11858-021-01238-1

Jankvist, U. T., & Niss, M. (2015). A framework for designing a research-based "maths counsellor" teacher programme. *Educational Studies in Mathematics, 90*(3), 259–284. https://doi.org/10.1007/s10649-015-9629-8

Jaworski, B., & Potari, D. (2021). Implementation of a developmental model of teachers' and didacticians' learning through inquiry: Design, operationalisation and outcomes. *ZDM - Mathematics Education*. https://doi.org/10.1007/s11858-021-01290-x

Karsenty, R. (2021). Implementing professional development programs for mathematics teachers at scale: What counts as success? *ZDM - Mathematics Education*. https://doi.org/10.1007/s11858-021-01250-5

Kieran, C., Krainer, K., & Shaughnessy, J. M. (2012). Linking research to practice: Teachers as key stakeholders in mathematics education research. In M. A. Clements, A. J. Bishop, C. Keitel, J. Kilpatrick, & F. K. S. Leung (Eds.), *Third international handbook of mathematics education.* Springer. https://doi.org/10.1007/978-1-4614-4684-2_12

Kontorovich, I., & Bartlett, P. (2021). Implementation of research on scriptwriting in an undergraduate mathematics course: A study of teacher-researcher collaboration. *ZDM - Mathematics Education*. https://doi.org/10.1007/s11858-021-01281-y

Krainer, K. (2021). Implementation as interaction of research, practice and policy. Considerations from the Austrian initiative IMST. *ZDM - Mathematics Education*, 53(5). https://doi.org/10.1007/s11858-021-01300-y

Kuzle, A., & Biehler, R. (2015). Examining mathematics mentor teachers' practices in professional development courses on teaching data analysis: Implications for mentor teachers' programs. *ZDM - Mathematics Education, 47*(1), 39–51. https://doi.org/10.1007/s11858-014-0663-2

Kwon, O. N., Bae, Y., & Oh, K. H. (2015). Design research on inquiry-based multivariable calculus: Focusing on students' argumentation and instructional design. *ZDM - Mathematics Education, 47*(6), 997–1011. https://doi.org/10.1007/s11858-015-0726-z

Lawler, J., Trevatt, P., Elliot, C., & Leary, A. (2019). Does the diabetes specialist nursing workforce impact the experiences and outcomes of people with diabetes? A hermeneutic review of the evidence. *Human Resources for Health, 17*, 65. https://doi.org/10.1186/s12960-019-0401-5

Levine, D. U., & Cooper, E. J. (1991). The change process and its implications in teaching thinking. In L. Idol & B. Jones (Eds.), *Educational values and cognitive instruction* (pp. 387–408). Routledge.

Maaß, K., & Artigue, M. (2013). Implementation of inquiry-based learning in day-to-day teaching: A synthesis. *ZDM - Mathematics Education, 45*(6), 779–795. https://doi.org/10.1007/s11858-013-0528-0

Maaß, K., & Doorman, M. (2013). A model for a widespread implementation of inquiry-based learning. *ZDM - Mathematics Education, 45*(6), 887–899. https://doi.org/10.1007/s11858-013-0505-7

Manouchehri, A., & Goodman, T. (2000). Implementing mathematics reform: The challenge within. *Educational Studies in Mathematics, 42*(1), 1–34. https://doi.org/10.1023/A:1004011522216

Mason, J. (1998). Researching from the inside in mathematics education. In A. Sierpinska & J. Kilpatrick (Eds.), *Mathematics education as a research domain: A search for identity* (pp. 357–377). Kluwer.

Mason, J. (2002). *Researching your own practice: The discipline of noticing*. Routledge.

Nilsen, P. (2015). Making sense of implementation theories, models and frameworks. *Implementation Science, 10*, 53. https://doi.org/10.1186/s13012-015-0242-0

Niss, M. (1996). Mathematics in society. In R. Biehler, R. W. Scholz, R. Sträßer, & B. Winkelmann (Eds.), *Didactics of mathematics as a scientific discipline* (pp. 367–378). Kluwer.

Niss, M. (1999). Aspects of the nature and state of research in mathematics education. *Educational Studies in Mathematics, 40*(1), 1–24. https://doi.org/10.1023/A:1003715913784

Obara, S., & Sloan, M. (2010). Classroom experiences with new curriculum materials during the implementation of performance standards in mathematics: A case study of teachers coping with change. *International Journal of Science and Mathematics Education, 8*(2), 349–372. https://doi.org/10.1007/s10763-009-9176-9

Pinto, A., & Koichu, B. (2021). Implementation of mathematics education research as crossing the boundary between disciplined inquiry and teacher inquiry. *ZDM - Mathematics Education*. https://doi.org/10.1007/s11858-021-01286-7

Polotskaia, E., & Savard, A. (2018). Using the relational paradigm: Effects on pupils' reasoning in solving additive word problems. *Research in Mathematics Education, 20*(1), 70–90. https://doi.org/10.1080/14794802.2018.1442740

Prediger, S., Fischer, C., Selter, C., & Schöber, C. (2019). Combining material- and community-based implementation strategies for scaling up: The case of supporting low-achieving middle school students. *Educational Studies in Mathematics, 102*(3), 361–378. https://doi.org/10.1007/s10649-018-9835-2

Prytz, J. (2021). When research met policy: A history of innovation and a complicated relationship in three Swedish development projects in mathematics education, 1960–2018. *ZDM - Mathematics Education*. https://doi.org/10.1007/s11858-021-01247-0

Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research, 75*(2), 211–246. https://doi.org/10.3102/00346543075002211

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*. https://doi.org/10.1007/s11858-021-01220-x

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.

Sierpinska, A., & Kilpatrick, J. (Eds.). (1998). *Mathematics education as a research domain: A search for identity*. Kluwer.

Smythe, E., & Spence, D. (2012). Re-viewing literature in hermeneutic research. *International Journal of Qualitative Methods, 11*(1), 12–25. https://doi.org/10.1177/160940691201100102

Star, S. L. (2010). This is not a boundary object: Reflections on the origin of a concept. *Science, Technology, and Human Values, 35*(5), 601–617. https://doi.org/10.1177/0162243910377624

Stylianides, G. J., & Stylianides, A. J. (2009). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education, 40*(3), 314–352. https://doi.org/10.5951/jresematheduc.40.3.0314

Sullivan, P., Jorgensen, R., Boaler, J., & Lerman, S. (2013). Transposing reform pedagogy into new contexts: Complex instruction in remote Australia. *Mathematics Education Research Journal, 25*(1), 173–184. https://doi.org/10.1007/s13394-013-0069-4

Tamborg, A. L. (2021). Improving mathematics teaching via digital platforms? Implementation processes seen through the lens of instrumental genesis. *ZDM - Mathematics Education*. https://doi.org/10.1007/s11858-021-01282-x

Valentine, N., Durning, S., Shanahan, E. M., & Schuwirth, L. (2021). Fairness in human judgement in assessment: A hermeneutic literature review and conceptual framework. *Advances in Health Sciences Education, 26*(2), 713–738. https://doi.org/10.1007/s10459-020-10002-1

Valoyes-Chávez, L., & Felmer, P. (2021). "She was probing me to see if I knew": Becoming a credible and confident PD facilitator. *ZDM - Mathematics Education*. https://doi.org/10.1007/s11858-021-01283-w

Venkat, H., & Graven, M. (2017). Changing teaching through a resources approach. In H. Venkat & M. Graven (Eds.), *Improving primary mathematics education, teaching and learning* (pp. 163–178). Palgrave Macmillan.

Wang, T. Y., Lin, F.L. & Yang, K.L. (2021). Success factors for a national problem-driven program aimed at enhancing affective performance in mathematics learning. *ZDM - Mathematics Education*, *53*(5). https://doi.org/10.1007/s11858-021-01285-8

Williams, S. R., & Leatham, K. R. (2017). Journal quality in mathematics education. *Journal for Research in Mathematics Education, 48*(4), 369–396. https://doi.org/10.5951/jresematheduc.48.4.0369

Wright, V. (2014). Frequencies as proportions: Using a teaching model based on Pirie and Kieren's model of mathematical understanding. *Mathematics Education Research Journal, 26*(1), 101–128. https://doi.org/10.1007/s13394-014-0118-7