A Delphi-inspired study on the concepts of large and small-scale projects in mathematics education research

Mario Sánchez Aguilar[1], Linda Marie Ahl[2], Uffe Thomas Jankvist[3] and Ola Helenius[4]

[1]Instituto Politécnico Nacional, CICATA Legaria, Mexico; mosanchez@ipn.mx

[2]Uppsala University, Department of Education, Sweden

[3]Aarhus University, Danish School of Education, Denmark

[4]University of Gothenburg, Sweden

We report on a study investigating how the notions of large-scale and small-scale projects are interpreted. Some of the methodological procedures of the Delphi technique were used, including posing a couple of scale-related questions to a group of 14 educational researchers who had experience in leading or investigating the large-scale implementation of educational innovations or reforms. The experts' responses were analyzed using Coburn's (2003) conceptualization of scale. The results show that these notions are generally interpreted from a perspective emphasizing quantitative aspects such as the number of teachers, schools, or districts reached by an implemented innovation or reform.

Keywords: Implementation research, large-scale project, small-scale project, conceptualization.

Introduction

Expressions such as "large-scale study" or "scaling-up" frequently appear in discussions and conceptualizations about the implementation of educational innovations (e.g., Century & Cassata, 2016; Coburn, 2003), including the field of mathematics education (e.g., Ahl et al., 2022; Heck et al., 2019; Roesken-Winter, Schüler, et al., 2015). Although the notion of "scale" is widely used in educational research, its meaning tends often to be handled tacitly, and there is diversity in how the term "scale" is used (Morel et al., 2019). The field of mathematics education research is not exempt from this lack of conceptual clarity in relation to the use of the notion of scale. With some exceptions (e.g., Krainer, 2015), notions such as "scaling-up", "at scale", or "large-scale" are used implicitly, giving rise to a variety of interpretations of these terms. In pursuit of conceptual clarity, we have undertaken a Delphi-inspired study that explores how the notions of "large-scale project" and "small-scale project" are understood in mathematics education research. The results of our study constitute a first attempt at how to characterize a "large-scale project" and a "small-scale project" in mathematics education. The following research question guides the study: *In relation to implementation research, how are the notions of a "large-scale project" and a "small-scale project" interpreted in mathematics education research?*

Although it may be impossible to answer this question in a definite manner, at least in this paper, an approximation is offered through an analysis of the opinions of experienced researchers. To analyze the experts' opinions, we adopt a classic conceptualization of the notion of "scale" proposed by Coburn (2003), which is informed by research on reform implementation.

Conceptual framework

Coburn's (2003) conceptualization of "scale" emerges as a response to the dominant perspective in educational research that tends to conceptualize studies at scale in a unidimensional way, that is, "involving solely or predominantly the expansion of numbers of schools reached by a given reform effort" (p. 3). Coburn argues that an adequate conceptualization of scale should go beyond merely spreading and considering all implementation challenges of reforms documented by decades of implementation research. Hence, Coburn proposes a multidimensional conceptualization of scale that not only focuses on the issue of spread but also considers other critical elements in implementing and adopting innovations, such as substantial changes in classrooms, endurance over time, and changes in authority. This conceptualization of scale consists of four interrelated dimensions: *depth*, *sustainability*, *spread*, and *shift in reform ownership*.

As part of the data analysis for this study—which will be explained in more detail in the method section—indicators were defined for each of these four dimensions. Next, the dimensions and the indicators defined for the data analysis are explained. The indicators are illustrated with mathematics education literature related to the notion of scale.

Depth. This dimension focuses on the quality and impact of reforms within individual classrooms or schools. In other words, it assesses whether there is a profound and consequential change that goes beyond surface structures and procedures and that has the potential to modify teachers' beliefs, social interaction norms, and underlying pedagogical principles.

Indicators. The depth indicators considered are utterances that refer to some kind of deep change within the classroom or the school system. For example:

- "To what extent can resources for mathematics classroom teaching, assessment, and professional development lead to changes in day-to-day classroom teaching?" (Maass et al., 2019b, p. 307)

Sustainability. This dimension refers to the possibility of reforms to persist over time despite political, financial, or organizational changes. Whether the reforms can endure and continue to impact even after the initial implementation can be measured in terms of sustainability. The dimension involves the critical conditions that must be present in schools and school districts to support and sustain classroom change over time.

Indicators. The indicators associated with this dimension are utterances that refer to the sustainability of the implementation of an innovation or to the conditions that must be present in a specific context so that the changes produced by an innovation or reform are lasting and sustained. For instance:

- "What comes after the professional development program is over? How can we sustain what was gained and how can the impact be scaled up?" (Tirosh et al., 2015, p. 153)

Spread. In this dimension, the focus is on spreading reforms or innovations to a more significant number of classrooms and schools. However, this dimension not only considers the possibility of spreading solely in terms of outward expansion to more and more schools; it also considers the spread of a reform or innovation within a classroom, a school, or a district.

Indicators. The spread indicators are those utterances that refer to an outward expansion of reform or innovation but also those that refer to a 'spreading within'. For instance:

- "This extends to around 40,000 teachers spread over the country's 290 municipalities and nearly 6,000 schools. The government has set aside SEK 650,000,000 (EUR 75,000,000) for the 4-year programme. This is one of several large-scale initiatives to improve mathematics teaching in Sweden" (Boesen et al., 2015, p. 129)

Shift in reform ownership. This dimension concerns the transfer of control and ownership of reforms from outside organizations to teachers, schools, or districts. It assesses whether the reforms are becoming embedded within the organizational culture and whether they are being led and sustained by those they directly impact.

Indicators. The indicators of this dimension are utterances that reference the presence of structures or mechanisms favoring ongoing learning about the reform or innovation, which provide funding and support for the implementation of the reform, or that reference how internal actors take responsibility for a continued spread of the reform. An example of such an indicator is:

- "Effective scaling requires teachers to obtain ownership of the reform and ultimately transfer the new approaches into an "internal reform" in their own school" (Potari et al., 2019, p. 419)

Method

We claim that the study reported in this article is a "Delphi-inspired" study because although it follows some of the methodological procedures of the Delphi technique, it does not strictly adhere to these. In the following, we describe the Delphi technique and the procedures from this technique in more detail. We also describe how the data analysis informed by Coburn (2003) was developed.

The Delphi technique and the procedures implemented

The Delphi technique is a method used to gather expert opinions or predictions on a specific topic. Its use as a research technique is not new in mathematics education (e.g., Manizade & Mason, 2011). It is a systematic, interactive forecasting process that utilizes a panel of experts to provide input through several rounds of questionnaires. The purpose of the Delphi technique is to achieve a consensus or to converge towards a common understanding, based on the experts' knowledge and experience (Barrett & Heale, 2020).

The Delphi technique typically involves the following methodological procedures:

1. Identifying and selecting a panel of experts in the relevant field.
2. Sending an initial questionnaire to the experts to gather their initial opinions and predictions.
3. Compiling and summarizing the results from the first round of questionnaires.
4. Sending a second questionnaire to the experts, incorporating the results of the first round and asking them to revise their opinions in light of the results.
5. Repeating this process until a consensus is reached or there is no further change in the opinions of the experts.

The current study followed Delphi technique's procedures 1, 2, and 3.

In the case of procedure 1, based on our experience and knowledge of the educational field and its literature, we identified and selected a group of 14 educational researchers from different regions of the world who had experience in leading or investigating the large-scale implementation of educational innovations and/or reforms. All educational researchers except for two are mathematics educators. However, these two exceptions are experts in scale studies, whose concepts are widely used in implementation research, including mathematics education research. All 14 experts were invited to participate in developing the second procedure of the Delphi technique.

For procedure 2, the 14 researchers were invited via email to answer two scale-related questions. All invited researchers agreed to participate in the study. They come from Europe, Latin America, and North America. Some decided to participate anonymously, while others consented to us using their names. It was explained to the researchers that the questions posed and their answers were part of the development of a review on implementation projects that included small-scale and large-scale projects (Ahl et al., 2023). However, deciding if an implementation project is small-scale or large-scale is difficult, and some may fall between the two categories. With this context as background, researchers were asked the following two questions:

- How do you define a large-scale project or a project at scale?
- Are “large scale” and “at scale” synonyms?

As part of procedure 3, the researchers’ email responses were collected in a shared document for analysis. Although the questions posed to the researchers involved various scale-related notions, in this paper, we only report on the interpretations related to the notions of large-scale and small-scale projects.

Data analysis

The analysis of the empirical data (the experts’ responses) was guided by the four dimensions that constitute the conceptualization of scale proposed by Coburn (2003). First, indicators for each of these dimensions were defined (see the “Conceptual framework” section). Afterward, the presence of the indicators in the explicit definitions or characterizations of the concepts of “large-scale project” and “small-scale project” stated by the experts were sought out. The absence or presence of these indicators in the empirical data were considered as markers of the dimensions of scale that are considered in the interpretations of a “large-scale project” and a “small-scale project” in mathematics education research.

Results

All of the surveyed experts explicitly referred to either the notion of a large-scale project or small-scale project or both—this, of course, was motivated by the fact that they were asked to define large-scale projects in one of the questions posed to them. The responses of the experts can be divided into two groups.

The first group of responses state that large-scale and small-scale projects are fuzzy concepts, difficult to define, or without a clear definition. Seven experts referred to these conceptual difficulties (e.g., “There is no absolute border between small and large scale”, Expert #3; “The only thing we could say about it was that different people treat it differently”, Expert #5; “There is no clear definition”, Expert #11; “The difference between small-scale and large-scale projects is quite challenging”, Expert #13).

The second group of responses refers to those that explicitly characterize these concepts. Nine experts provided these types of responses, which were analyzed in search of the indicators of the dimensions of the Coburn (2003) framework. The results show that all explicit references to the notions of large-scale and small-scale contain *spread indicators*. For example, “Projects that focus on instructional improvement at scale can be quite small scale (i.e., require limited funding and involve just a few people).”, Expert #1; “Of course, no one would deny that mathematics education research studies involving thousands of students or teachers are large-scale studies”, Expert #6; “We consider our PD project as a large-scale project as we offer PD for facilitators who themselves conduct PD for teachers”, Expert #11. However, a couple of the researchers interviewed warned us that the issue of size in large-scale projects is relative. For instance, Expert #3 states: “Above 100 teachers is not small scale any more, but what is really large might also depend on the size of the country”.

Indicators of *depth* and *sustainability* were only identified in the responses of one of the interviewed mathematics education researchers. In one of his answers, the expert refers to how large-scale projects should “permeate the system”, which is interpreted as a reference to a significant change in the system in which an innovation is implemented: “‘Large scale’ means ‘big’ (often costly as well) but something may be big without necessarily permeating a system.”, Expert #14. When referring to large-scale projects, this researcher mentions “stability”, which is interpreted as a reference to the potential of the implemented projects to remain stable over time, i.e., sustainability: “There are teaching support programs called ‘math circles’. There are many such organizations—see https://mathcircles.org/. The number of organizations, and their stability, says that they are large scale”, Expert #14.

Nevertheless, in four formulations of large-scale or small-scale projects, we identified features that are not included in Coburn’s (2003) conceptualization of scale. These features refer to differences in methodological conditions and tools and notions such as representativeness and generalization: “There is no doubt for me, that whatever are the distinctions made, studies at different scales differentiate in terms of methodological conditions and constraints, which has impact on their possible outcomes”, Expert #6; “One criterion could be if you can do quantitative, statistical analysis on the group? Versus small scale if only qualitative analysis makes sense?”, Expert #12; “Large scale is rather vague, I may suggest; is it all the population or large enough so that then a generalization is accepted? Actually, it is the question of generalization which is important and must be the criterion”, Expert #8.

Discussion

The responses from seven out of the 14 experts indicate that it is challenging to distinguish between the notions of large-scale and small-scale projects. This fact emphasizes the relevance of developing a study that looks for conceptual clarity on these notions.

Regarding the question that guides this study, “*In relation to implementation research, how are the notions of a ‘large-scale project’ and a ‘small-scale project’ interpreted in mathematics education research?*”, a first approximation to the answer would be that these notions are generally interpreted from a perspective of scale that emphasizes the spread dimension. The dominant perspective consists of quantitatively interpreting these notions, prioritizing the number of teachers, schools, or districts that are reached by implementing a reform or innovation. Although there is evidence of mathematics educators' interpretations that include dimensions of depth and sustainability, this perspective is not dominant.

The prevalence of this scale interpretation among mathematics educators, which tends to emphasize the spread dimension, is corroborated when contrasting their conceptualizations of scale with those of the two educational researchers (non-mathematics educators) who participated in the study. The educational researchers' conceptualizations of scale also include the dimension of spread, but the notion as a whole is portrayed as an ultimate state fully realizing all targeted scaling dimensions:

> There is a difference between a scaling project and a project that is "at scale." Saying "at scale" suggests that the project has accomplished the goal of having fully realized all targeted dimensions of scaling., Expert #4

> We define "spread" as the process of an innovation becoming more widespread. Similar terms include diffusion and scaling. We define "scale" as the end-state of spread—that is how something is used once it has become widespread. So, for us, scale is a goal or end-state., Expert #9

Moreover, some experts' large-scale and small-scale formulations include features not encompassed in Coburn's (2003) conceptualization of scale, such as differences in methodological conditions and tools and the notions of representativeness and generalization. These characterizations point to the need for a conceptualization of scale that not only focuses on aspects of impact, reach, stability, and independence of the implemented project, but which also considers the differences in approaches and tools that enable the study of its effects, as well as the potential for generalization to other populations and contexts.

The study reported in this paper has continued to evolve, complementing the interpretations identified in this Delphi-inspired study with specialized literature on implementing innovations at scale in mathematics education (see Aguilar et al., 2023 for details). In particular, we have reviewed 22 articles from two special issues related to implementing and scaling innovations, published in *Educational Studies in Mathematics* (Maass et al., 2019) and *ZDM — Mathematics Education* (Roesken-Winter et al., 2015). The articles were read and analyzed, extracting data on scale definitions and explaining what was referred to when talking about small scale, large scale, scaling, and at scale. In the elaboration of this paper, we suggest considering factors such as contact, material, and organizational factors to differentiate between small- and large-scale implementation projects. Additionally, we introduce the concept of medium-scale implementation projects, which can bridge the gap between small- and large-scale projects. We also provide examples of scenarios in which medium-scale projects can be conceptualized.

The insights gained from the elaboration of this research can inform future implementation research in mathematics education by providing a characterization for understanding and discussing scale and scaling—and this not only in terms of quantitative measures but also regarding the qualitative features of the innovations in question.

Acknowledgment

This paper is part of grant 2022-04811 under the *Swedish Research Council*.

References

Aguilar, M. S., Ahl, L. M., Jankvist, U. T., & Helenius, O. (2023). Towards characterization of scale and scaling in implementation research within mathematics education. *Implementation and Replication Studies in Mathematics Education*, *3*(1), 99–122. https://doi.org/10.1163/26670127-bja10012

Ahl, L. M., Aguilar, M. S., Misfeldt, M., Koichu, B., & Jankvist, U. T. (2022). Recent discussions on implementation research at CERME12 [Editorial]. *Implementation and Replication Studies in Mathematics Education*, *2*(1), 1–19. https://doi.org/10.1163/26670127-02012015

Ahl, L. M., Helenius, O., Aguilar, M. S., Jankvist, U. T., Misfeldt, M., & Prytz, J. (2023). *A theoretical framework for analyzing factors of influence for the success or failure of implementation projects* [Manuscript in preparation]. Uppsala University; University of Gothenburg; Instituto Politécnico Nacional; Aarhus University; University of Copenhagen.

Barrett, D., & Heale, R. (2020). What are Delphi studies? *Evidence-Based Nursing*, *23*(3), 68–69. https://doi.org/10.1136/ebnurs-2020-103303

Boesen, J., Helenius, O., & Johansson, B. (2015). National-scale professional development in Sweden: Theory, policy, practice. *ZDM – Mathematics Education*, *47*(1), 129–141. https://doi.org/10.1007/s11858-014-0653-4

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

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, *32*(6), 3–12. https://doi.org/10.3102/0013189X032006003

Heck, D. J., Plumley, C. L., Stylianou, D. A., Smith, A. A., & Moffett, G. (2019). Scaling up innovative learning in mathematics: Exploring the effect of different professional development approaches on teacher knowledge, beliefs, and instructional practice. *Educational Studies in Mathematics*, *102*(3), 319–342. https://doi.org/10.1007/s10649-019-09895-6

Krainer, K. (2015). Reflections on the increasing relevance of large-scale professional development. *ZDM – Mathematics Education*, *47*(1), 143–151. https://doi.org/10.1007/s11858-015-0674-7

Maass, K., Cobb, P., Krainer, K., & Potari, D. (Eds.). (2019a). Different ways to large scale implementation of innovative teaching approaches [Special issue]. *Educational Studies in Mathematics*, *102*(3).

Maass, K., Cobb, P., Krainer, K., & Potari, D. (2019b). Different ways to implement innovative teaching approaches at scale. *Educational Studies in Mathematics*, *102*(3), 303–318. https://doi.org/10.1007/s10649-019-09920-8

Manizade, A. G., & Mason, M. M. (2011). Using Delphi methodology to design assessments of teachers' pedagogical content knowledge. *Educational Studies in Mathematics*, *76*(2), 183–207. https://doi.org/10.1007/s10649-010-9276-z

Middleton, J. A., Cai, J., & Hwang, S. (Eds.). (2015). *Large-scale studies in mathematics education*. Springer. https://doi.org/10.1007/978-3-319-07716-1

Morel, R. P., Coburn, C., Catterson, A. K., & Higgs, J. (2019). The multiple meanings of scale: Implications for researchers and practitioners. *Educational Researcher*, *48*(6), 369–377. https://doi.org/10.3102/0013189X19860531

Potari, D., Psycharis, G., Sakonidis, C., & Zachariades, T. (2019). Collaborative design of a reform-oriented mathematics curriculum: contradictions and boundaries across teaching, research, and policy. *Educational Studies in Mathematics*, *102*(3), 417–434. https://doi.org/10.1007/s10649-018-9834-3

Roesken-Winter, B., Hoyles, C., & Blömeke, S. (Eds.). (2015). Evidence-based CPD: Scaling up sustainable interventions [Special issue]. *ZDM – Mathematics Education*, *47*(1).

Roesken-Winter, B., Schüler, S., Stahnke, R., & Blömeke, S. (2015). Effective CPD on a large scale: Examining the development of multipliers. *ZDM – Mathematics Education*, *47*(1), 13–25. https://doi.org/10.1007/s11858-014-0644-5

Tirosh, D., Tsamir, P., & Levenson, E. (2015). Fundamental issues concerning the sustainment and scaling up of professional development programs. *ZDM – Mathematics Education*, *47*(1), 153–159. https://doi.org/10.1007/s11858-015-0665-8