



HAL
open science

Innovation and the problem of measuring effects in large scale professional development projects

John Prytz, Linda Marie Ahl, Uffe Thomas Jankvist

► To cite this version:

John Prytz, Linda Marie Ahl, Uffe Thomas Jankvist. Innovation and the problem of measuring effects in large scale professional development projects. Thirteenth Congress of the European Society for Research in Mathematics Education (CERME13), Alfréd Rényi Institute of Mathematics; Eötvös Loránd University of Budapest, Jul 2023, Budapest, Hungary. hal-04407691

HAL Id: hal-04407691

<https://hal.science/hal-04407691>

Submitted on 21 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Innovation and the problem of measuring effects in large scale professional development projects

John Prytz¹, Linda Marie Ahl¹ and Uffe Thomas Jankvist²

¹Uppsala University, Department of Education, Sweden; johan.prytz@edu.uu.se;

²Aarhus University, Danish School of Education, Denmark

This paper concerns the relation between the research fields of implementation research (IR) and professional development (PD). Our main interest is in the concept of innovation and measurement of effects in large scale development projects. The case for our discussion is the Swedish large-scale development project Boost for Mathematics (BM). A project designed according to theories from PD research. Our tentative conclusion, in an IR perspective, is that it is difficult to replicate the BM project. This despite the fact there are six studies that measure the effects of the project. We argue that replication requires better knowledge about the innovations as possible causes of effects. We suggest further research in this direction.

Keywords: Innovation, effects, measurement, development project, replication.

Introduction

The study presented in this paper is part of the project *Implementation research as an emerging field of mathematics education*. The project's aim is to create an empirically based theoretical framework for implementation research (IR) in the field of mathematics education research, in particular for research on large-scale development projects. Our project examines existing theories in implementation research and which parts of these are applicable to implementation projects related to mathematics education. To verify this, we test the theories and related concepts that we find relevant through comparisons of five selected cases. Historical comparisons are made of three Swedish projects: Boost for Mathematics (BM) (*Matematiklyftet*) (2013–2016); PUMP (*Processanalyser av Undervisning i Matematik/Psykologisk*, Process analysis of Teaching in Mathematics/Psycholinguistics) (1974–1977); and New Mathematics (1961–1968). Also, contemporary comparisons are made between Boost for Mathematics and MIST (Middle-school Mathematics and the Institutional Setting of Teaching) (2007–) and TRIUMPHS (TRansforming Instruction in Undergraduate Mathematics via Primary Historical Sources) (2015–), the two latter are projects from the USA.

In this paper we present an initial and tentative study of the relation between theories from IR and professional development (PD) research. These two fields are entangled since large scale professional development projects, concerning mathematics education, are enterprises where innovations often are implemented; we know this from our recent research overview (Ahl et al., 2022). By studying, through an IR lens, a large-scale development project based on PD theory we seek to understand in what way central elements of IR theory, in comparison to PD theory, can provide new insights; this is our purpose. As a case for our discussion, we take the Swedish BM project.

The BM was an in-service training program for mathematics teachers in Sweden, covering school years 1 to 12. The project ran between 2013 and 2016. 76% of all mathematics teachers joined the project. It is well known the BM was designed according to PD theories (Boesen et al., 2015; Prytz,

2021). A core design idea was peer learning among teachers. Prytz (2021) concluded that teachers rather than researchers were supposed to complete the innovations. Teachers were given much freedom in relation to the training material. Another feature of the BM is that its impact on teachers' behaviour and students' learning have been measured in several studies (Grönqvist et al., 2021; Lindvall, 2017; Lindvall et al. 2022; Lindvall et al., 2018; Ramböl, 2016; Österholm et al., 2016).

However, from an IR perspective the idea to measure impact or effect of a project where teachers are given much freedom is problematic. In particular when causes of the effects are considered. In this paper we develop a theoretical discussion on why this is problematic, with the BM as our case. Our guiding questions for this discussion are:

1. What type of project was the BM in an IR-perspective (*adaption* or *fidelity* centred)?
2. What were the innovations of the BM and in what way were they innovations?
3. What effects have been measured and how have they been explained in studies of the BM?

As to the paper's disposition, this introduction is followed by a section on theory. In that section, we present a theoretical difference between the research fields of IR and PD. The difference concerns the concept of innovation. We also present two concepts used to answer question 1: *fidelity* to an innovation and *adaption* of an innovation. Thereafter, in a section about analytical approach, there is a brief description of how the guiding questions are answered. Then an initial analysis is presented, where the main point is that there are much to do in order map all innovations of the BM. In a final section we discuss, with the BM as example, how theories from the fields of IR and PD are related and how that relation can be problematic. We also suggest further empirical studies of the BM.

IR theoretical constructs

In the field of IR, the concept of innovation is central. It is the thing to be implemented. For instance, Century and Cassata (2016), in their overview on IR, defines the field as follows:

... we offer a working definition of implementation research as 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)

In contrast, the role of innovations is not as salient in the PD field. Take for instance the following definition in an overview on PD research by Darling-Hammond et al. (2017):

In this review, we define effective professional development as structured professional learning that results in changes to teacher knowledge and practices, and improvements in student learning outcomes. We conceptualise professional learning as a product of both externally provided and job-embedded activities that increase teachers' knowledge and help them change their instructional practice in ways that support student learning. Thus, formal PD represents a subset of the range of experiences that may result in professional learning. (p. 2)

Here, PD is understood as structured professional learning. And if PD changes the knowledge and practice of teachers or student learning, it is effective. Teachers are also supposed to engage in specific activities and be exposed to experiences. No explicit formulations about innovations are made, though. Still, in the seven properties which, according to Darling-Hammond et al. (2017),

define effective PD programs, the innovation concept appears in an implicit way. More precisely, teachers are supposed to do or focus on certain things, which can mean that teachers meet something new. For instance, a PD program should focus on discipline-specific pedagogy, teachers should share ideas, experts can give support about evidence for best practice, and there should be time to reflect on best practice. From this we conclude, when the concept of innovation is part of a PD research, it is an implicit part and far from its main focus. Nevertheless, this means there is a potential overlap between IR and PD when it comes to innovations. However, as soon as the impact of PD-projects are measured, the role and nature of the innovations become relevant. In particular, when one enters discussions about causal relations. Our basic thought is this: if there is an effect, there should be a cause. In an educational context, it is thus relevant to understand this cause as something new (an innovation) going into the teachers' mind or into their teaching.

If we now turn to the IR approach to measurement of effects, it clearly involves the innovation concept, but also the fidelity concept. Fidelity in this context means that teachers understand and use the innovation as planned by its proponents. In order to measure the effects of an innovation, we need to know if teachers have used the innovation with fidelity. Following Century and Cassata's (2016) overview on IR, fidelity is a matter of degrees. Perfect fidelity is not possible to achieve in practice and is neither the only ideal to pursue. The teaching situation is a complex one, and we must assume that teachers adapt an innovation to local circumstances. Coburn (2003) even points out that to achieve sustainable implementation of an innovation over time (a quality and desirable effect), adaptation can be necessary. However, this does not make the concepts of innovation and fidelity irrelevant when considering effects. It is still important to identify the core components of an innovation, to identify influential factors in the implementation process, and to know in what respect the enactment of the innovation differs from the theoretical ideal (Century & Cassata, 2016). Even in cases where the sought effect or outcome is sustainability, which can require adaptation of the given innovation, it remains relevant to compare how the innovation is used in practice with the original plan. Because if the core components of the innovation have been lost, it is not a matter of sustainability. For a longer discussion of the fidelity concept, see Jankvist et al. (2022).

Analytical approach

In order to answer the first guiding question, we draw on results from our previous studies which in part concern the BM. The original contribution of this paper is to apply certain IR concepts (fidelity and adaptation) not used in the previous studies. For the second and third guiding questions, our material consists in reports and research articles about the BM. We then seek to distinguish what the planned innovation of the BM were, but also which effects of the BM have been measured, and how these have been explained. For the second question, we have also used two online search engines to distinguish in what respect the content of the BM material was innovations for teachers. Or if the teachers can have encountered the essential BM content through other channels.

One search engine is *Retriever* (<https://www.retrievergroup.com/sv/product-mediarkivet>). *Retriever* has the biggest media archive in the Nordic countries covering printed and online editorial media as well as radio and television since the 1980s. We have restricted our search to media in the Swedish language. The other engine is the article search engine of the teacher journal *Nämnamn*

(<https://ncm.gu.se/namnarenopenaccess-sok>). The journal is dedicated to mathematics education and is targeted towards teachers in mathematics. Its search engine only handles *Nämnamnaren* material. From *Retriever*, we wanted to see when important elements of the BM got general media coverage in Sweden. From the search engine for *Nämnamnaren*, we wanted to see when important elements of BM were introduced more specifically to teachers. In this way, we can start to discern what were actually the new or innovative important elements of the BM and in what way.

Initial analysis of the Boost for Mathematics project

The BM was in certain respects a top-down project. In the two decades before the BM, Swedish results in international large-scale assessments had decreased very much, which received much media attention. Politicians were forced to act. One of the attempts to counter this development was the BM. But it was not the only attempt, for instance, a new mathematics curriculum was issued in 2011, that is two years before the launch of the BM. During the preparations of the BM, a plausible cause of the poor results was identified: teachers were too seldom leading the teaching and students were too often working alone without supervision and feedback from the teachers. This was reflected in the general goals of BM: to achieve a changed teaching culture and a changed in-service training culture. A new teaching culture meant that teachers should have more teaching methods at their disposal and reflect better about teaching decisions (Prytz, 2021). So, initiated by the national government and organised by the National School Agency (*Skolverket*), the BM was a top-down project.

However, the top-down character did not entail that fidelity was central in the design of the BM. Indeed, there were certain and official goals, but they were quite open; teachers should get more teaching methods (not one) and develop their ability to reflect and make own decisions (Prytz, 2021). This openness was further emphasised by the project's basic design principle: peer learning among teachers. This entailed in-service training in the schools, managed by selected senior teachers at the schools. Before the training began, the selected teachers received an eight-day course in mathematics education. These courses were given by university departments specialised in mathematics education (Prytz, 2021). Given these circumstances, Prytz (2021) concluded that teachers were given the role of completing the innovations, while the material for the training constituted a context of innovation.

As mentioned in the introduction, it is well-known that the BM was designed according to PD theories (Boesen et al., 2015; Prytz, 2021), which included the peer learning design. For example, in official reports and in decision documents concerning BM, there were explicit references to an international overview on PD research. However, Prytz (2021) pointed out that this free role of the teachers was also a product of a school governance policy that favoured decentralisation and teacher autonomy.

In the light of basic IR theory, we therefore argue that *adaption* rather than *fidelity* was central in the BM. Teachers were given freedom to adapt innovations to local circumstances; they were not given specific guidelines they had to abide by.

So, what about innovations in the BM? To understand what the planned innovations were, the training material is an important source. But, it is a vast material and, in this paper, we only scratch the surface. For the training, there were so-called modules for the topics of arithmetic, geometry, algebra, and functions. There were also modules for the topics problem solving, digitalization, and language in mathematics. For all topics, there were sets of modules dedicated to the year-spans 1–3, 4–6, 7–9,

and 10–12. All sets of modules also covered four themes: abilities (competencies), formative assessment, interaction, and socio-mathematical norms. The concrete content of the modules was films, audio clips, and texts with instructions and questions for the peer learning sessions. The modules also contained scientific articles: international overviews of each topic and the four themes; and specific articles concerning the same topics and themes in a Swedish context. The modules were distributed via the National School Agency website (Prytz, 2021).

Participation in the program meant that each group of teachers in a school had to take two sets of modules. The modules in arithmetic were mandatory, while the second was optional, but one set had to be chosen. The modules had been created by researchers working in university departments specialised in mathematics education. In certain respects, the modules *per se*, in combination with the peer learning idea, can be seen as an innovation of BM. As far as we can see, this type or form of PD was something new in Sweden, at least at this scale. And the innovation was supposed to meet the goals of changing the in-service training culture and the teaching culture.

As to the content of the modules, there were also innovations. At the time of the beginning of the BM, two of the modules' four themes (formative assessment and socio-mathematical norms) were quite new to teachers. We used the two media search engines to see how often the words “formative assessment” and “socio-mathematical norms” were used in Swedish media in general and in the teacher journal *Nämna* in the period of 1980 to 2022. The Swedish words are “formativ bedömning” and “sociomatematiska normer”. As to media in general, “formative assessments” began to appear much more frequently after 2009, 11 publications that year, increasing fast to a peak with 266 publications in 2015. In total we got 1617 publications. In *Nämna* we got 37 publications, almost all in 2009 or after, except for two publications. This picture is confirmed by an overview by Hirsch and Lindberg (2015) about Swedish research on formative assessment. According to this overview the number of Swedish PhD theses about formative assessment increased after 2010. In contrast, the expression “sociomathematical norm” was very rare, it hardly appeared in media in general (7 publications in total) or in *Nämna* (4 publications in total). All 11 publications appeared after 2010, with a concentration to the period 2011–2015.

Regarding the concepts of “abilities” and “interaction”, it is not meaningful to search for those, since they have a wider meaning and can appear in many contexts. However, we know they appeared already in the 1994 mathematics curriculum; both had a prominent role (Prytz, 2015). We also know that the curriculum messages about abilities and interaction were complex and teachers had difficulties understanding them (Boesen et al, 2014). In the 2011 curriculum reform, a similar type of message was pushed once again (Prytz, 2015). Thus, the BM can be considered a new reform tool in that implementation process. The innovation was then not the ideational content, but the form in which the content was presented to teachers. The fact that a message or a theme was implemented via a curriculum reform can mean that teachers could get further information about it through for instance textbooks, which also were affected by the curriculum reform. This can be compared to how socio-mathematical norms was barely mentioned outside the BM-modules. Our point is that the four themes as innovations appear to have been different in nature (ideational or form); and teachers' possibilities to learn about them through other media may have differed. But, further studies of for

instance textbooks and curriculum documents are required to corroborate this point. Further studies of the BM material are also needed to better understand how the four themes were concretized.

Now to the measuring and explanations of effects of the BM. Two major studies measuring a changed teacher culture have been made: Ramböl (2016) and Österholm et al. (2016). Both studies were planned parts of the BM. In brief, both studies only involved teacher which participated in the BM; the Ramböl (2016) analysis was focused on organisational aspects, while Österholm et al. (2016) was also focused on ideational aspects and included the modules' four themes (abilities, formative assessment, interaction, and socio-mathematical norms) in the analysis. Both studies indicated that teachers perceived a new culture and a change in teaching. The studies also indicate that the actual teaching had changed. Concerning their explanations, Ramböl (2016) identified the BM in-service training model as a cause of the positive effects. Österholm et al. (2016) could not identify any particular organisational factor causing positive effects; but they concluded that participation in the BM seems to have been the important factor. They also identify that the teaching culture had changed, in varying degrees, along the four themes. However, the possibility that the teachers might have received relevant information about some themes from elsewhere was not considered in the analysis.

Another set of studies consider the BM's effect on student learning: Grönqvist et al. (2021), Lindvall (2017), Lindvall et al. (2022), and Lindvall et al. (2018). Grönqvist et al. (2021) was a planned part of the BM. A very brief and general summary of these four studies is that they seek to find out if and in what way the BM and its PD design had an impact on student learning. All four had a control-experiment-group-design. Students whose teachers had participated in the BM were in the experiment group, while students whose teachers had not participated in the BM were the control group. Different types of test results were used. All four studies discerned positive effects on student learning in some school years, but far from all. To varying degrees the four studies considered the content of the BM and how it affected the teachers and student learning. Lindvall et al. (2018) was the most specific and focused only on the arithmetic modules, providing also a detailed content analysis of these modules. A common feature of the other three studies is that the content of the BM is treated as one unity. As to the four BM themes, they are mentioned in all four studies, but they constitute one unit rather than several different innovations. Consequently, the measured effects were caused by this unit. How different innovations played out was not studied.

Conclusions

To measure and assess effects of a large-scale PD program such as the BM is relevant, in certain perspectives. The fact that the BM reached 76% of all Swedish teachers in mathematics and cost nearly 100 million Euros of tax-money are good reasons; the relevant question for tax-payers, politicians, and teachers is then: was it worth it? But also, from a PD research perspective it is relevant, since the basic design of the BM was based on PD theory, more precisely a principle about peer learning among teachers. And the six studies, referred to in this paper, measuring different effects of the BM, do give relevant answers to how the peer learning design functioned and in what way. Still, as we have indicated, from an IR perspective, the measurements of the BM are problematic. In IR, innovation is essential when effects are measured. You need to know what it is, its core

components, and you need to know to what degree teachers had fidelity to the innovation. Our initial analysis suggests a number of potential problems in the BM with respect to this:

1. Adaptation was favoured over fidelity.
2. Different types of innovations (ideational or form) were at play at the same time.
3. Innovations could to varying degrees have been mediated through other channels.

This makes it difficult to identify which innovations or types of innovations should be considered a cause of the measured effects. This also makes it difficult to discern the degree of fidelity. Add to that our analysis, so far, does not concern innovations connected to mathematical content. Thus, there might be more innovations that can complicate the situation further.

But what makes up a problem from an IR point of view does not need to be a problem within PD research. In this paper, it is also noted that IR and PD theories put different emphasis on the innovation concept; in IR theory it is explicitly positioned as essential, while in PD theory it is more implicitly in the background. Hence, it is not surprising that we in IR see problems in a project based on PD theory. And, as pointed out above, it makes sense in a PD perspective to measure the BM in the way it has been measured. So why should PD people care? The key concept in our answer is *replication*. If you want to carry out a Boost for Mathematics 2—or at least draw on experiences from the BM in later large-scale projects—it is relevant to retain those innovations that functioned and omit those that did not function well. It then becomes relevant to analyse further what the BM innovations actually were and determine their nature. See for example Koichu et al. (2021) for a discussion on different natures of innovations (material-centered vs. interaction-centered).

We therefore intend to make further studies into the BM materials (modules) and make comparisons with for instance textbooks and curriculum documents, which we presented at CERME. More precisely, it is possible and relevant to compare BM modules with curriculum reform innovations. We know for instance that the new national curriculum in 2011 had stronger focus on communication, but also problem solving. To such comparisons we can add comparisons of BM modules with popular textbooks published both before and after the 2011 curriculum reform.

Acknowledgements

This paper is part of grant 2020-04090 under the *Swedish Research Council*.

References

- Ahl, L. M., Aguilar, M. S., Jankvist, U. T., Misfeldt, M., & Prytz, J. (2022). Professional development as a means for implementing mathematics education innovations: results from a systematic review. In Mattsson, L. Häggström, J., Carlsen, M., Kilhamn, C., Palmér, H., Perez, M., & Pettersson, K. (Eds.) *The relation between mathematics education research and teachers' professional development. Proceedings of MADIF 13*. (pp. 57–67). Swedish Society for Research in Mathematics Education.
- Boesen, J., Helenius, O., & Johansson, B. (2015). National-scale professional development in Sweden : theory, policy, practice. *ZDM - the International Journal on 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>
- Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute. <https://doi.org/10.54300/122.311>
- Grönqvist, E., Rosenqvist, O., & Öckert, B. (2021). *Lyfter Matematiklyftet matematikkunskaperna?* [Does Boost for Mathematics boost knowledge in mathematics?] IFAU – Institutet för arbetsmarknads- och utbildningspolitisk utvärdering.
- Hirsh, Å., & Lindberg, V. (2015). *Formativ bedömning på 2000-talet – en översikt av svensk och internationell forskning* [Formative assessment in the 21st century – an overview of Swedish and international research] [Electronic resource]. Vetenskapsrådet.
- Jankvist, U. T., Aguilar, M. S., Misfeldt, M., & Koichu, B. (2022). How about Fidelity? *Implementation and Replication Studies in Mathematics Education*, 2(2), 131–148.
<https://doi.org/10.1163/26670127-02012020>
- Koichu, B., Aguilar, M. S., & Misfeldt, M. (2021). Implementation-related research in mathematics education: the search for identity. *ZDM–Mathematics Education*, 53(5), 975–989.
<https://doi.org/10.1007/s11858-021-01302-w>
- Lindvall, J. (2017). Two large-scale professional development programs for mathematics teachers and their impact on student achievement. *International Journal of Science and Mathematics Education*, 15(7), 1281–1301. <https://doi.org/10.1007/s10763-016-9750-x>
- Lindvall, J., Helenius, O., & Wiberg, M. (2018). Critical features of professional development programs: Comparing content focus and impact of two large-scale programs. *Teaching and Teacher Education*, 70, 121–131. <https://doi.org/10.1016/j.tate.2017.11.013>
- Lindvall, J., Helenius, O., Eriksson, K., & Ryve, A. (2022). Impact and design of a national-scale professional development program for mathematics teachers. *Scandinavian Journal of Educational Research*, 66(5), 744–759. <https://doi.org/10.1080/00313831.2021.1910563>
- Prytz, J. (2015). Swedish mathematics curricula, 1850–2014. An overview. In *Dig where you stand 3: Proceedings of the Third International Conference on the History of Mathematics Education*. (pp. 309–326). Uppsala University.
- 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*, 53(5), 1035–1046. <https://doi.org/10.1007/s11858-021-01247-0>
- Ramböl (2016). *Slututvärdering av matematiklyftet 2013–2016* [Final evaluation of the Boost for Mathematics 2013–2016]. Ramböl.
- Österholm, M., Bergqvist, T., Liljekvist, Y., & van Bommel, J. (2016). *Utvärdering av Matematiklyftets resultat: slutrapport* [Evaluation of the results of the Boost for Mathematics: final report]. Umeå University.