LUMAT

Special Issue: Educational Design Research

Vol 7 No 3
2019

Contents

Editors 2

Editorial 3
Johannes Pernaa and Veli-Matti Vesterinen

Teachers as actors in an educational design research: What is behind the generalized formula? 6
Helén Sterner

Design and evaluation of practice-oriented materials fostering students' development of problem-solving competence: The case of working backward strategy 28
Ana Kuzle

Characteristics of teacher knowledge produced by pre-service mathematics teachers: the case of open-ended problem-based learning 55
Jani Hannula

Pedagogical experiments with MathCheck in university teaching 84
Terhi Kaarakka, Kirsi Helkala, Antti Valmari and Marjukka Joutsenlahti

Towards student-centred solutions and pedagogical innovations in science education through co-design approach within design-based research 113
Maija Aksela

A systematic review of educational design research in Finnish doctoral dissertations on mathematics, science, and technology education 140
Daranee Lehtonen, Anne Jyrkiäinen and Jorma Joutsenlahti

Editors

Johannes Pernaa

Johannes Pernaa works at the Department of Chemistry in University of Helsinki. As a university lecturer of chemistry education his main research area is to develop chemistry teacher education through research. Johannes' research focus is information and commucation technology in chemistry education, which includes a lot of learning environment development through educational design research. Johannes has over decade of experience in design-based research including doctoral dissertation from 2011 and multiple research articles. His has also worked as the editor of book called "Design-based Research in Education" (published in 2013, Finnish title "Kehittämis-tutkimus opetusalalla").

Veli-Matti Vesterinen

Veli-Matti Vesterinen works as a university lecturer at the Department of Chemistry, University of Turku. As an educational design researcher he is especially interested in various forms of collaboration and co-design in science and sustainability education. His other research interests include nature of science, inquiry learning, science communication as well as students' and teachers' role as sociopolitical actors.

Editorial

Johannes Pernaa[1] and Veli-Matti Vesterinen[2]

[1] The Unit of Chemistry Teacher Education, Department of Chemistry, University of Helsinki, Finland
[2] Department of Chemistry, University of Turku, Finland

During the past few decades, several interconnected research traditions have paid more and more attention to the process of educational design. Educational design research and other design-oriented methods seek complex educational problems through systematic, iterative, and continuing process of design, development, and evaluation of educational practices. This special issue presents six articles including research on educational design research methodology as well as research utilizing educational design research methods.

Correspondence: johannes.pernaa@helsinki.fi; veli-matti.vesterinen@utu.fi

DOI: https://doi.org/10.31129/LUMAT.7.3.442

Educational design research

Educational design research (EDR), also known as design-based research (DBR), is a methodological approach that enables systematic research-oriented praxis. EDR can be used for developing all kind of educational artefacts like e.g. learning materials, courses and software (Pernaa & Aksela, 2013).

EDR studies have been conducted since the early nineties (Brown, 1992; Collins, 1992). Nowadays it is a well-known and widely used approach for educational research (Anderson & Shattuck, 2012). Scholarly interest on the EDR can be seen, for example, in the growing number of handbooks, guides and special issues in educational research journals focusing on it (e.g. Barab & Squire, 2004; Kelly, 2003; Pernaa, 2013; Plomp & Nieveen, 2010; Sandoval & Bell, 2004).

This special issue presents six studies, which either utilise or discuss the use of EDR in the context of mathematics, science and technology education research. They provide concrete examples on how to use the approach in the development of materials and approaches as well as to develop our understanding of the EDR methology.

In the first paper, Helén Sterner discusses about the teachers' role in educational design research. She explores the opportunities and challenges teachers encounter when they participate in educational design research projects. The study shows, that teachers' pariticipation in all the phases of educational design research helps to focus the project to the instructional practices of the participating teachers. Thus such participation can play a central part in interweaving research and practice.

In the second paper, Ana Kuzle reports a design research project on the development of practice-oriented materials for supporting students' problem-solving competences were developed. Problem-solving through working backward strategy was selected as the focus for the learning environment design. Kuzle has selected this context, because according to earlier research literature, it has been found difficult for students to learn and use.

Jani Hannula reports a design research project in which a course for mathematics education has been developed. The aim of the course was to strengthen connections between university-level mathematics and school mathematics, which is an important issue in mathematics teacher education. Hannula presents the course design process and a case study carried out in the designed course. The study focuses on pre-service mathematics teacher knowledge produced during an open-ended problem-based learning task.

Terhi Kaarakka, Kirsi Helkala, Antti Valmari and Marjukka Joutsenlahti introduce MathCheck – an online application supporting mathematics learning by giving learners feedback on their math solutions. MathCheck is a research-based software designed through constructivism learning theory. They study MathCheck's effect on learning via five pedagogical experiments.

In the fifth paper, Maija Aksela focuses on studying the collaboration occurring in a diverse multi-stakeholder educational design research project. The aim of her case study is to demonstrate how a co-design approach can be used within such projects. Possibilities and challenges of co-design approach are analysed through a large framework, where a large design community is developing several student-based solutions and pedagogical innovations simultaneously

The last paper is a systematic review of Finnish doctoral dissertations applying research-based design methods in the context of mathematics, science and technology education research. In their review of 21 recent dissertations, the authors Daranee Lehtonen, Anne Jyrkiäinen and Jorma Joutsenlahti provide an overview on how educational design research has been used and developed in Finland.

References

Anderson, T., & Shattuck, J. (2012). Design-Based Research: A Decade of Progress in Education Research? *Educational Researcher*, *41*(1), 16–25. https://doi.org/10.3102/0013189X11428813

Barab, S., & Squire, K. (2004). Design-Based Research: Putting a Stake in the Ground. *Journal of the Learning Sciences*, *13*(1), 1–14. https://doi.org/10.1207/s15327809jls1301_1

Brown, A. (1992). Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings. *The Journal of Learning Sciences*, *2*(2), 141–178.

Collins, A. (1992). Toward a Design Science of Education. In *NATO ASI Series. New Directions in Educational Technology* (pp. 15–22). https://doi.org/10.1007/978-3-642-77750-9_2

Kelly, A. E. (2003). Theme Issue: The Role of Design in Educational Research. *Educational Researcher*, *32*(1), 3–4. https://doi.org/10.3102/0013189X032001003

Pernaa, J. (Ed.). (2013). *Kehittämistutkimus opetusalalla* (Eng. Design-based Research in Education). Jyväskylä: PS-kustannus.

Pernaa, J., & Aksela, M. (2013). Model-Based Design Research: A Practical Method for Educational Innovations. *Advances in Business-Related Scientific Research Journal*, *4*(1), 71–83.

Plomp, T., & Nieveen, N. M. (Eds.). (2010). *An introduction to educational design research: Proceedings of the seminar conducted at the East China Normal University, Shanghai (PR China), November 23-26, 2007.* Enschede: SLO.

Sandoval, W. A., & Bell, P. (2004). Design-Based Research Methods for Studying Learning in Context: Introduction. *Educational Psychologist*, *39*(4), 199–201. https://doi.org/10.1207/s15326985ep3904_1

Teachers as actors in an educational design research: What is behind the generalized formula?

Helén Sterner

Linnaeus University and Dalarna University, Sweden, Faculty of Mathematics Education

Educational design research provides opportunities for both the theoretical understanding and practical explanations of teaching. In educational design research, mathematics teachers' learning is essential. However, research shows that little consideration is given to teachers and the participation of teachers throughout the entire design process as well as in continued learning. With this in mind, educational teacher-focused design research was used to explore the challenges teachers face, and the opportunities teachers are given when they participate as actors in all the phases of educational design research - designing, teaching, and refining theoretical concepts within the teaching. In this study, the mathematics focus of the design research was generalizations in patterns with Design Principles as the theoretical frame. The results show that the participation of teachers in all the phases of a design process is central for the teachers' learning. Moreover, challenges that the teachers encounter in the classroom provide opportunities and consequences for the continued design process and lead to changes in the teachers' understanding of generalizations. The results also indicate that functional thinking and linear equations contributed to both the teachers' and students' learning about generalizations in patterns.

Keywords
design principles,
elementary school,
grades 1-6,
functional thinking,
generalization,
patterns,
teachers

Correspondence
hse@du.se

DOI
https://doi.org/10.31129/LUMAT.7.3.403

1 Introduction

Educational design research provides opportunities for both theoretical and pragmatic orientations and creates opportunities for interventionist methodology and collaboration between researchers and teachers. Several previous studies give creative examples of students learning (e.g., Stephan & Akyuz, 2012) or give examples of how research teams to design, implement, and refine teaching to improve classroom practice with teachers and develop the theoretical understanding of different classroom phenomena (e.g., Stephan, 2015; Stephan & Akyuz, 2012). The results from previous design research studies have advocated theoretical solutions for practices. However, implementing theoretical concepts into practice could lead to certain problems; for example, after the design research has been carried out, the teacher is still responsible for the continued teaching.

Mathematics education researchers have highlighted a lack of classroom-based interventions that focus on teaching and teacher learning and also the lack of design-based studies promoting investigation in the classroom. This gap requires a shift in research focus from the learners to the teachers (Mariotti, Durand-Guerrier & Stylianides, 2018). One way to approach teaching is to focus on teachers and the participation of teachers in order to understand how they communicate when designing tasks and teach theoretical concepts within a specific mathematical perspective. To this intention, studying the participation of teachers and the meaning that becomes negotiated (Wenger, 1998) in a specific mathematical content within a process may be central. When teachers play an active role, and the object of study is teachers' meaning-making of a specific algebraic perspective, the results can be a new, presumable, and challenging way to think about mathematics teaching in elementary school (Grade 1-6). In Grades 1-6 in Sweden (grundskola årskurs 1-6), it is not common to teach functional thinking in relation to patterns, thus making it particularly interesting to follow a design process that includes functional thinking because it deviates from a more traditional way of teaching. This also makes a reasonable argument for why the teachers must negotiate the meaning of this mathematical content.

The aims of this article are to shed light on the process when teachers participate as actors in educational design research focusing on generalizations in patterns and to elaborate on the consequences these challenges have for teachers in their learning process. What challenges do teachers face, and what opportunities arise when teachers negotiate the meaning of generalization in patterns in educational design research?

2 Design research and the content in the intervention

Educational design research provides for theoretical contributions and understanding of classroom phenomena as well as possibilities to develop classroom practice. A variety of educational design research and professional development programs are documented, named, and designed in a number of common ways, for example, developmental research (Gravemeijer & van Eerde, 2009), classroom design studies (Stephan & Akyuz, 2012) and professional development design studies (Cobb & McClain 2001). Traditionally, educational design research involves the intervention and recurring cyclical processes of designing, teaching, and refining teaching experiments in a specific educational domain.

2.1 Focus in educational design research

Educational design studies are conducted in diverse ways. Some research focuses on students' learning and learning activities (Stephan & Akyuz, 2012), while others focus on teachers' learning in a design research process (Cobb & McClain, 2001). Few design studies and interventions focus on teachers and teachers' learning in all phases of the design process - designing, teaching, and refining. In one study, classroom design research focusing on students' learning, also involves three practices that the teachers learned and identified when participating in an intervention: collaboration with other teachers, collaboration with the researcher, and learning from their reflections on their teaching (Stephan, 2015). Another study place teachers in the role of co-designer in the design process. Here the authors describe the teacher as a crucial person who goes from being an implementer to a co-designer (Konrad & Bakker 2018).

Various methodologies can be implemented in educational design research. In this study, a teacher-focused classroom design study was used, which can be regarded as a classroom design study that focuses on teachers' learning instead of students' learning. Classroom design studies are characterized by researchers often collaborating with several mathematics teachers in a particular mathematical domain. Certainly, limitations in educational design studies have been identified, which Cobb, Jackson, and Dunlap Sharpe (2017) describe as "the lack of attention to the instructional practices of the teacher in the study" (p. 228). This quotation serves to further highlight the focus of the teacher and her practice.

Teachers' learning in educational design research

Educational design research is intended to produce relevant and usable theoretical knowledge that can be used in practice. However, research has yielded further questions in teacher development programs, such us, what have teachers actually learned and experienced in these programs (Sztajn, Borko & Smith 2017)? Questions of what teachers have learned led to a greater understanding of teachers' difficulties to use and implement theoretical concepts from research in the teaching, which also falls in line with Cobb et al. (2017), among others. This is yet another reason why the teachers' learning is central in all phases in the design process.

Design Principles

In educational design research, a theory is used to frame the content in the intervention. Normally, design research has recurring cyclical processes and uses, for example, hypothetical learning trajectories for a specific educational issue. These hypothetical learning trajectories are continuously revised and unfold in a local instruction theory in the predetermined educational area (Gravemeijer & van Eerde, 2009). However, in this study, recurring cyclical processes are still used, but instead of hypothetical learning trajectories, Design Principles (DPs) identified from previous research are used as a theoretical frame (Greeno, 2006; McKenney & Reeves, 2012). The DPs give the intervention a theoretical grounding, guide the intervention, and provide opportunities to understand a particular educational phenomenon (Greeno, 2006). The purpose was to develop a conceptual understanding of generalizations in patterns and functional thinking. In turn, this led to identifying DPs from previous research that has been essential for teaching generalization in patterns and functional thinking. The DPs can be considered similar to goals for teaching that will give theoretical ideas of the content, which students then have the opportunity to learn and develop.

2.2 The mathematical content in the intervention

One specific perspective of algebra and algebraic thinking has been chosen for this study – patterns in arithmetic sequences. Algebraic thinking in Grades 1–6 can be likened to mathematical reasoning or algebraic reasoning, and algebraic thinking is sometimes equated with early algebra (Blanton, Stephens, Knuth, Garidner, Isler & Kim, 2015). In this study, algebraic thinking also means giving students opportunities to generalize and justify relationships among quantities. Therefore, generalization, patterns, and functional thinking are central in this article.

Generalization

Generalization is expressed as a key aspect of algebraic reasoning (e.g. Kieran 2004). The concepts of generalization are used in both everyday life and in mathematical contexts with a variety of explanations and meanings. According to Dörlfler (1991), generalizations in mathematics can be defined as both “an object and means of thinking and communication” (p. 63). In school mathematics, the object could be interpreted as the generalizations expressed in conventional symbols, and the thinking and communication could be likened to communicating and reasoning with generalizations expressed through a variety of representations.

The central elements of reasoning in school mathematics include activities such as exploration, conjecture, and justification. These concepts can be seen as activities in a cyclical process, such as the reasoning and proof cycle (NCTM 2008). Both conjecture and justification involve activities of logical thinking and general statements that make it possible to view and express mathematical structures, thus extending generalizations (Blanton, Stephens, et. al. 2015). However, research indicates that teachers and students' have difficulties understanding and working with generalizations (Stylianides & Silver, 2009). Working with generalizations and pattern identification in teaching is embedded with difficulties, such as understanding the differences and the nuances that exist in pattern identification "in which cases can we trust a pattern without a need for further examination of particular cases" (p. 249). This quotation relates to empirical generalization and theoretical generalization – empirical generalization can be described as generalizing from one situation to another, and theoretical generalizing is generalizing to an abstraction. Dörfler (1991) considers that some form of symbolic description, for example, letters, geometric illustrations or verbal stories, are needed to make theoretical generalizations. Teachers draw on different resources when teaching generalizing in mathematics, and according to Dörfler (1991), empirical generalization and theoretical generalization can be seen as problematic in mathematics teaching. One reason may be that mathematics as a science uses theoretical generalizing, whereas several general conceptions in both school mathematics and in real life use empirical generalizations as explanatory models. Dörfler (1991) highlights the importance of being aware of the mutual relationship between empirical and theoretical generalizations.

Patterns and functional thinking

One way to introduce generalization in elementary school is through teaching patterns and pattern identification. When working with patterns in general, opportunities for algebraic thinking develop (Blanton, Stephens, et. al. 2015; Mulligan, Cavanagh & Keanan-Brown, 2012; Mulligan, Mitchelmore, English & Crevensten, 2013). According to Kaput (2008), algebraic thinking involves two core aspects – making and expressing generalizations in symbol systems and reasoning with symbolic forms.

Studies suggest that neither mathematics curricula nor mathematics teaching may prepare students enough to transform from concrete arithmetic reasoning to abstract algebraic reasoning (Blanton, Stephens, et al., 2015; Kieran 2004). Therefore, the process of first-arithmetic-then-algebra is regarded by some as controversial, as the teaching in elementary school often includes both algebraic and arithmetic concepts (e.g., Blanton, Stephens, et al. 2015). One aspect of algebraic reasoning is functional thinking, which includes for example identifying a recursive pattern and describing the pattern in words, identifying a covariational relationship and describing the relationship in words, identifying the meaning of a variable used to represent a varying quantity, and constructing a coordinate graph (Blanton & Kaput, 2004; Blanton, Stephens et al., 2015; Carraher & Schliemann, 2015). According to Beckmann and Izsák (2015), understanding the relationship between two (or more) varying quantitates is central when working with functions and functional thinking. Beckmann and Izsák (2015) highlight the concept of the variable-parts perspective to express a proportional relationship and facilitate the identifying of functional relationships between two covarying quantities. This variable-parts perspective could explain how qualities can remain fixed even when quantities vary in a proportional relationship. Elementary students often meet unknown quantity with a missing fixed value (Blanton, Levi, Crites & Dougherty, 2011); however, studies have shown that elementary students represent the functional relationship by using variable notation to represent the generalization (e.g., Blanton, Brizuela, Gardiner, Sawrey, & Newman-Owens 2015; Blanton, Stephens, et al. 2015; Schliemann, Carraher, & Brizuela, 2007). What is interesting in these studies is that the young students chose to use variable notation to denote the generalizations rather than tell a story with natural language to express the generalization.

One way to think about algebra in elementary school is expressed by Caspi and Sfard (2012), who relate school algebra and generalization to a meta-discourse of arithmetic, including numerical patterns, relationships, and quantitates. Caspi and Sfard (2012) point to the value of teaching algebra in students' daily practices in close relation to the students' understanding of informal arithmetic. The authors emphasize the importance of teaching and present parts of the algebraic thinking in several mathematical issues in elementary school before students meet formal algebra in upper secondary school. Caspi and Sfard (2012) emphasize the importance of acknowledging the close relationship between formal algebra and the students' understanding of arithmetic thus, "helping the students closing the gap between their spontaneous meta-arithmetic and the formal algebra taught in school" (p. 21). This quotation in particular, points to unresolved issues in teaching algebra in elementary school, namely teachers' challenges to give students the opportunities to understand formal algebra. The spontaneous meta-arithmetic and the formal algebra, as well as Dörfler's (1991) interpretation of a mutual relationship between empirical generalization and theoretical generalization, fall somewhat in line with what has emerged in research –that young students prefer to use variable notation rather than natural language to express generalizations.

2.3 Algebra and functional thinking in the Swedish context

Mathematics teaching in Sweden is characterized by textbook use, as a large part of the teaching is guided by a textbook (Johansson 2006). In Swedish curriculum materials (National Agency of Education, 2017) for elementary school, mathematical reasoning is one competency that students should be given the opportunity to develop through a variety of mathematical content. The concepts of generalizing are not mentioned in the Swedish national curriculum materials for elementary school; instead, patterns and proportional relationships are conceptualized as two separate mathematical themes in the materials (National Agency of Education, 2017). These themes are often treated separately in the Swedish classroom and textbooks, and the topic of proportional relationships in elementary school Grades 1–6 (Swedish *grundskola årskurs 1–6*) is often used to explain the concepts of "double" and "a half" and is rarely used as functional thinking in the lower grades. Functional thinking and the equation for linear function are first described in Grades 7–9 in curriculum materials (National Agency of Education 2017). Carraher, Schliemann, and Schwartz (2008) have also pointed out the lack of functional relationships in elementary school: "It is nothing short of remarkable that the topic of functions is absent from early mathematics curricula" (p. 265).

3 Overview of the study

The study outlined in this article is part of a project entitled *Mathematical reasoning in algebra in elementary school – The role of the teacher in an educational design research process*. In this article, one aspect of mathematical reasoning is central – generalizations – and generalizations in a specific algebraic perspective, namely, patterns and functional thinking. In this study, conducting educational design research using teachers as actors and involving teachers in the entire design process was a way to explore the challenges that teachers encountered when designing, teaching, and refining theoretical concepts related to generalizations in patterns.

3.1 Theoretical framing

When focusing on teachers and their communication relating to generalizations in teaching patterns, theoretical frames that visualize teachers' learning were necessary. From the perspective of Communities of Practice, (CoP), learning is explained as a process in which participants engage in social practice (Wenger, 1998). Therefore, learning can be interpreted as participants' changed participation and changed communication about a mutual engagement. The process involving participants' meaning-making and reification is an active dynamic process, and in a process, the participants' shared repertoire changes. In this case, the participants negotiate the meaning of the content in the DPs, which means the teachers in the study communicate and relate generalizations about patterns and functional thinking to their teaching.

Based on a part of Wenger's (1998) learning theory, the idea of learning during boundary crossing was central in this study. In this article, the intervention will be seen as a boundary encounter in the teachers' learning. In line with CoP and Wenger, various boundary encounters can happen in a CoP, for example, a conversation between two participants from two different communities involving some sort of boundary. Another way to enrich the boundary encounter in the community is to visit another practice to gain some exposure. Boundary encounters consist of boundary objects. Wenger (1998) describes these boundary objects as new and unknown objects for the participants' in the community, for example, artifacts, terms, or concepts. Thus, the boundary objects could be a way to organize the communication and learning in a CoP. In this case, the intervention can be comparable to a boundary encounter – the participants meet theoretical concepts about generalizing in pattern and functional thinking. These theoretical concepts of the DPs' are comparable to the boundary objects as unknown concepts introduced in the teachers' learning. In the design process, the teachers participate as actors and have possibilities to negotiate the meaning of the DPs related to their teaching.

To summarize, Wenger's (1998) notions of *boundary crossing, boundary encounter,* and *boundary objects* are used to explore and exemplify the challenges and opportunities that teachers face in a design process when they negotiate the meaning of generalizations in patterns. For this purpose, Wenger's analytical concepts of boundary-crossing, and boundary objects are applied as well as the negotiation of meaning and the participants' changed repertoire about generalizations in patterns.

3.2 Methodology

Conducting a teacher-focused classroom design research using teachers as actors throughout the design process was a way to explore the challenges that teachers encounter when designing, teaching, and refining theoretical concepts related to generalizations in patterns. The purpose of the intervention is to support the conceptual understanding of generalizations in patterns and functional thinking as well as to support teachers' learning in the process. However, the main purpose of this article is to visualize the challenges and opportunities that occur in educational design research. Therefore, two perspectives are central in this study – that of the teachers as they participate as actors throughout the entire process and that of a study situated in the entire design process. The red field in Figure 1 illustrates where the study is situated. The three phases – which symbolize designing, teaching and refining – overlap and are inevitably intertwined.

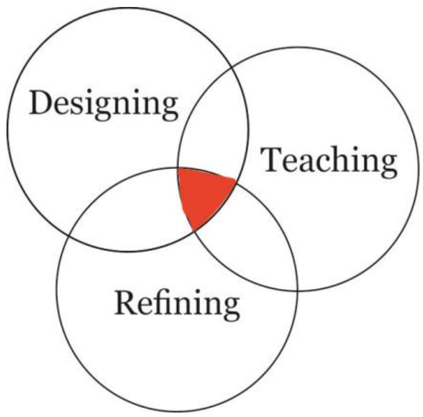


Figure 1. The red field represents where this study is situated.

Teachers as participating actors throughout the entire design process

In addition to the author, three mathematics teachers, one from Grade 1 and two from Grade 6, participated and collaborated in three recurring design cycles. The three teachers were selected from a previous study on teachers' meaning-making about mathematical reasoning and mathematical communication (Sterner, 2015). As a result, the participants and their community can already be seen as an incorporated CoP, according to Wenger (1998). Wenger explains that a CoP includes the participants having a shared way of engaging and doing things together, for example, a quick setup of a problem to be discussed and sustained mutual relationships that may be harmonious or conflictual.

The study took place over a period of nine months (Fig. 2) and three recurring design cycles were carried out. Figure 2 shows a timeline of the entire design study, which illustrates each individual teacher's process in the design study. The three half circles illustrate each of the three teachers' teaching, and the line underneath illustrates the designing and refining phases. This article will focus solely on designing and refining phases. Therefore, the teaching will be used to exemplify challenges that arise from the teaching and what opportunities and consequences these challenges will have for the continued design process.

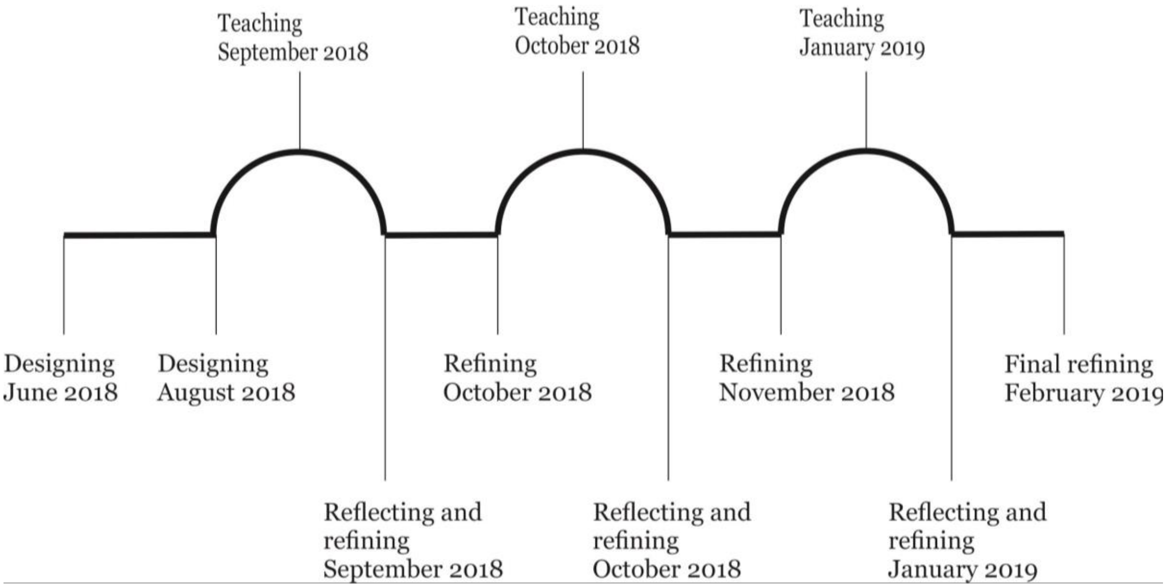


Figure 2. Timeline of the entire study.

Teacher-focused classroom design research

To take advantage of what has emerged in previous research on generalizations and patterns, DPs were chosen to focus on the mathematical content in the intervention. DPs are formulated based on previous research; however, the teaching is traditionally done in a Swedish context, which the intervention needed to take into account. These DPs frame and guide the intervention as well as give theoretical ideas of the content that the students have opportunities to learn and develop in the teaching. The two DPs used in this study were created as a goal for teaching:

DP1: The students should be given opportunities to identify a pattern, structure the pattern and generalize the pattern (Mulligan et al., 2013; Stylianides & Silver, 2009).

DP2: The students should be given opportunities to work with algebraic reasoning, including functional thinking and proportional relationship and determining relations between two or more varying quantities (Beckman & Izak, 2015; Blanton, Stephens, et al., 2015; Blanton & Kaput, 2004).

The site and the participants

The participating teachers work at three different schools in a mid-size town in Sweden. Three recurring cycles were carried out, and five meetings with teachers were carried out – two meetings to design activities before teaching and one session after completing the design research (Fig. 2). The researcher (the author) worked closely and collaborated with the teachers in the design and refining phases and acted as an observer in the teaching. After teaching, the teacher and the author spent 20–30 minutes reflecting and refining.

Each common meeting was two and a half hours long, and all the participants participated in each meeting. The purpose of these common meetings was to create and develop an understanding of the theoretical concepts in the DPs as well as design and refine the teaching that included these theoretical concepts. The individual reflection directly after the teaching was conducted in order to get the teachers' immediate thoughts and ideas in connection with what had just happened in the classroom. These reflections were used in the forthcoming common refining phase.

Data collection and the selected data

The data collection included video recordings from various parts of the design process: five common meetings from the design and refining phases and fifteen lessons consisting of twelve from Grade 6 and three from Grade 1. Other data material includes copies of student work, field notes from the classroom observations, and tape recordings of reflections with the teachers directly after teaching.

The empirical data analyzed in this study are primarily the three teachers' communication about generalizations in patterns and the actions they took related to generalizations. Teaching sequences from Grade 6 were used to exemplify activities and tasks related to the DPs. These sequences also influence the ongoing discussions in the common refining phase. The selected data were based on empirical data grounded on what challenges and opportunities teachers meet when they negotiate the meaning of generalizations in patterns in a design research project as well as in teaching.

Data analysis

In this study, the DPs are seen as a theoretical framework for the intervention and as goals for teaching. In the analysis, Wenger's (1998) theoretical frame of boundary objects revealed the teachers' meaning-making and understanding of generalizations in patterns based on the content of the DPs. Wenger's frame makes it possible to make visible how the teachers' shared repertoire of DPs changed in the teachers' discussions during the process as well as in the teaching. Focusing on boundary objects also made visible how the DPs were transformed from theoretical concepts to activities in the classroom and for the teaching and furthermore to refining the activities and the teaching.

Ethical consideration

The intention was not to generalize the results but rather to exemplify the challenges and opportunities of an educational teacher-focused classroom design research and give an understanding of what those can accomplish. Thus, the actual number of teachers is not relevant to this study. However, ethical considerations have to be made throughout the entire research process, from the first planning of the study to the last report (Goodchild, 2011). Both external and internal issues (Floyd & Arthur, 2012) have been important in this study. Internal issues include taking into consideration the teachers' and the students' participation as well as the chosen content in the design process. Therefore, it was important that I ask myself certain questions, for example, in what way can theoretical concepts from the mathematics research field and the competencies from mathematics teachers' practice interact? And in what way can our different competencies and experiences be visible and operated in the process? And what mathematical content is relevant and possible in this intervention? The external ethical issues regulated for research provided by the Swedish Research Council (2017) were followed, and both guardians and students gave consent for participation. The students were given verbal information about the intervention, and both the students and the students' guardians were given written information about the study and had approved participation in line with the ethical guidelines.

The context in which the study took place

The observed teachers' teaching was similar both in terms of structure and the teaching situation, although some differences naturally arose. To demonstrate the context of this study, Irma's teaching serves as an example, as it was broadly representative of the approach adopted by all three teachers, particularly the two Grade 6 teachers. Irma began the sequence of teaching patterns with multiple approaches when introducing patterns, particularly those to do with growing arithmetic sequences. She used various teaching methods and a variety of materials to illustrate patterns, for example, matches, cubes, and pictures. In Irma's teaching, several representations were revealed to depict the patterns; for example, pictures, tables, or students' stories.

Two mathematical tasks that were taken from other studies were particularly interesting in the intervention. One task was used to exemplify direct proportionality. All the teachers worked with a task concerning a certain number of dogs and the corresponding number of tails, ears, and legs to illustrate the use of independent and dependent variables. The students had to represent the number of tails, ears, and legs in different ways, for example, pictures, tables, diagrams, and graphs. The works of Blanton and Kaput (2004) and Mulligan et al. (2013) inspired this task, which was used to understand direct proportionality.

The other especially interesting task in the design process was a pattern illustrating "the cube train" (see Figure 3). The cubes represented a train with an engine and one or more train cars. The red cubes illustrated the engine, and the green cubes illustrated the first train car (Fig. 3). This task, in which the cubes illustrated a growing arithmetic sequence, was inspired by Beckman (2018): "The first train is made from a 4-cube engine and one 5-cube train car. Each subsequent train is made by adding one more 5-cube train car" (p. 415).

Figure 3. Image of the "the cube train", which represents a growing pattern in an arithmetic sequence (Beckman 2018, p. 415).

This cube train task created opportunities to talk about differences in direct proportionality. For example, in the previously mentioned task, the number of dogs and legs is represented by $y = 4x$. The cube train task offered possibilities to talk about and uncover the m –value or the 0th position (Beckman 2018).

The two tasks, the dogs and the train, were used to give students opportunities to use different representations to explain and find the generalizations. The tasks were chosen to clarify the m −value and shed light on the slope of the graph. Furthermore, the tasks created opportunities to talk about independent and dependent variables, as well as how the quantities corresponded. The overall aim of working with these tasks was to give opportunities for the students to work with and develop an understanding of generalization by working with patterns in arithmetic sequences.

4 Results

The results are presented based on the challenges and opportunities that arose in the teachers' meaning-making about generalizations in the design process. When the teachers worked as actors in the design process and negotiated meaning about generalization in patterns, five themes of challenges were revealed in the analysis: working with the unknown, understanding what DPs mean in the classroom, questioning taken-for-granted views, the need for mathematical language and finally starting to teach with the generalizations. The challenges were proven to lead to opportunities for developing teaching. These themes will be presented in the teachers' negotiation of meaning about generalizations in patterns in relation to their teaching.

4.1 Working with the unknown

The intervention and the DPs were found to be challenging in several ways. In the initial stage of the discussions in the design phase, the teachers clearly stated that they did not want to work with functional thinking or proportional relationships when teaching patterns. The teachers questioned the relevance between the two DPs and could not see the importance of working with two "totally different" contents, patterns, and functional thinking. They asked several times, "Why are we using linear equations together with patterns? ... we don't understand the reason." The teachers negotiated the meaning and tried to understand the theoretical concepts in DP1. One of the challenges during the process turned out to be working against the unknown, as in the boundary objects. These boundary objects (DP1 and DP2) turned out to create both tension and opportunities in the teachers' discussions to develop something that was not previously known in their teaching. Working with functional thinking in relation to patterns was a new and previously unknown way to think about mathematical content patterns and generalizations.

4.2 What do DPs mean for the classroom?

The next stage of the discussion in the design phase before the teaching consisted of teachers' meaning-making about the DP1. The teachers negotiated the meaning of each of the concepts in the DP1 and asked questions such as, what does it mean to identify? What does it mean to structure? And what does it mean to generalize a pattern?" The teachers' discussions focused on practical teaching ideas related to the DP1 – identifying, structuring and generalizing patterns. In the teachers' initial discussion, the three theoretical concepts were discussed in close relation to the teaching practice. In particular, the discussion focused on different representations in relation to identifying, structuring and generalizing. In the teachers' meaning-making, identifying a pattern was often associated with students' stories. Structuring patterns were often associated with students' pictures, tables, or conversations when the students used different materials, such as matches or cubes, and these materials were put into piles illustrating the pattern related to the training task (Fig. 3). In the teachers' negotiation of meaning, similar questions and assertions arose: "Is it possible to structure a pattern and equate it as a generalization?" and "Every time we added one more cube train car, we got 5 more cubes, and the 4-cube engine only happens once. It's the same is for the whole train." With regard to the teachers' discussion, in the initial stage, the three theoretical concepts in DP1 were understood as three different representations; later, they negotiated the meaning and talked about generalization as something that could be represented in different ways. Nevertheless, there was resistance to discussing the content in DP2; however, opportunities to discuss DP2 arose when the teachers talked about the task with the train. The teachers wanted to design similar tasks with an equally clear m –value.

4.3 Questioning taken-for-granted views

It became visible in the teaching that generalizing patterns could be represented in different ways, like identification in a story or a structured table as well as in general formula. The teachers talked about the students and their use of different representations to finally arrive at the general formula. In the end, the teachers were all in agreement that generalizing a pattern with a general formula would be the final, optimal understanding for students. In other words, a generalization was equated to a general formula. This statement was agreed with by the teachers until they tried to challenge the students' understanding and ask for justification and explanation of the general formula.

One of the tasks the teachers returned to several times was the growing pattern illustrated with a train: “made from a 4-cube engine and one 5-cube train car” (Fig. 3). In the teaching, the teachers gave the students possibilities to talk about their empirical stories and the patterns in the arithmetic sequence, for example, “Every sequence is increased by five cubes” or “The number of train cars five times and add four.” From this and similar stories, many Grade 6 students could use the language and generalize as well as write a general formula. Some of the students needed a table to structure the information before they see and could express the growing pattern. The teachers employed practical ideas from the discussions in which they had previously participated, for example, those from earlier discussions in which the teachers had agreed to use variable notations and a general formula that would represent the best generalizations and the final state of the activity.

In teaching, the teachers faced unexpected reflections from the students. Irma tried to challenge the students and ask them to explain and justify what $5x$ in this general formula $y = 5x + 4$ signified and symbolized. At that moment, few students could justify and explain the generalized formula, and most of the students did not have the words to talk about, what is behind the generalized formula. Later in the forthcoming reflection in the design process, the teachers discussed what they experienced during the teaching. They talked about the frustrating and eye-opening moments in their teaching. As teaching with patterns often includes several representations with justifications, they had previously taken-for-granted that the general formula with variable notation was the final representation and, therefore, no explanation would be needed for the formula.

4.4 Need for mathematical language

A change in the teachers’ communication arose, as the discussions changed to curiosity and thinking about opportunities to use functional thinking to help the students understand and explain the general formula. The meaning-making in their discussions led back to functional thinking, the DP2, and the task about the number of dogs relating to their number of tails, ears, and legs. The teachers also designed other tasks, such as, “What about other animals, like spiders with eight legs?” In the design and reflection phases, the teachers’ meaning-making about generalizations changed from seeing the generalized formula as the final representation to focusing on the generalized formula as one representation and functional thinking as a possibility to understand and talk about generalizations.

Working with linear functions gave the students as well as the teachers possibilities to talk differently about the generalizations and understand what is behind the generalized formula. Teaching linear equations and functional thinking was new and unknown for the teachers regarding teaching generalizations in patterns and arithmetic sequences. The teachers recognized new challenges, such as several unknown words that are necessary for working with functional thinking and proportional relationships, for example, "coordinates", "independent and dependent variables", "graphs," and "linear equations". This new unknown world brought with it a negotiation about the meaning of these worlds and a need for mathematical language in both the teachers' discussions and in the classroom.

At the end of the process, some students, as well as the teachers, said, "I can see the generalized formula in a graph." The teachers discussed the importance of having tasks with a clear $m -$ value. Here the teachers often returned to the exercise with the train, where the engine could be seen as a clear illustration of the $m -$ value.

4.5 Start to teach with the generalizations

In the last refining phase, the teachers discussed the students' co-variational thinking and how they use natural language to describe how the quantities corresponded. The teachers discussed the importance of the $m -$ value, the constant of proportionality (k), and the slope of the graph. They discussed how some students had the ability to see, understand, and describe the general formula with natural language. Therefore, the teachers wanted to design teaching and activities where students start with the general formula. The students relied on using natural language to talk about the slope, the independent and dependent variable, and the $m -$ value, and then used different representations to show the generalization.

The teachers designed different tasks and activities to make it possible for the students to use different representations to justify the generalized formula. These activities helped the teachers talk about different patterns in relation to generalizations. The teachers used the general formula given in the aforementioned train task:

> We know this is a general formula $y = 5x + 4$ to describe the growing pattern in the cube train task. Can we illustrate the formula in another way? What does the start-value mean ($m-$ value)? What are the differences between these equations: $y = 4x$, $y = 4x + 2$, $y = 6x + 2$, and $y = 6x + 4$?

The students justified the difference, for example, between $y = 4x$ and $y = 4x + 2$ and talked about the $m-$ value. The teachers discussed the students' competence in reasoning about two generalized quantities that are related, in that the ratio of one quantity to the other is invariant. One of the teachers said, "Nowadays, some students recognize a pattern of direct matches when they encounter a general formula, or a see a graph."

To summarize, this case visualizes the importance of the involvement of teachers in all phases – the teachers are faced with challenges in the classroom that have consequences for the conversation in the design and refining phase. It seems reasonable to infer that the teachers learn new things about generalizations; however, the teachers changed and talked differently about generalizations in patterns as well as changed and developed their teaching in generalizations. Therefore, functional thinking supports the teachers as well as the students when talking about the generalizations.

5 Discussion

The results of my study give insight into what can be gained from a teacher-focused classroom design research. The results show teachers involved in a complex process of meaning-making in the process of understanding teaching and learning generalizations in patterns in algebra.

In the study, the teachers participating as actors in educational design research faced various challenges in the process. These challenges were sometimes met with resistance; however, they gave rise to consequences that appear to drive the teachers' process of change and development. Therefore, meeting these challenges seemed to be a prerequisite for opening up opportunities; for example, certain challenges were revealed when teachers designed activities and teaching generalizations in patterns in arithmetic sequences. The teachers realized that the students had difficulties in justifying a generalized formula. In the design process, the teachers also became aware that functional thinking and the linear equation could create opportunities to talk about what is "behind" a generalized formula. The teachers found that teaching the concept of functional thinking and the linear equation in close relationship with patterns could facilitate students' understanding of generalizations. Previous studies also support these findings (e.g., Blanton & Kaput, 2004; Blanton, Brizuela, et al. 2015; Blanton, Stephens, et al., 2015). In this particular case, the teachers refer to functional thinking as something that creates opportunities to discuss and justify the generalized formula, and this led to a change in the teachers' awareness of generalizing. Generalization, as a concept beyond the more generalized formula and variable notation, is related to what has been found in previous studies – that elementary students use variable notation rather than natural language to express a generalization (Blanton, Brizuela, et al. 2015; Blanton, Stephens, et al., 2015). This falls in line with both Caspi and Sfard (2012), who claim the importance of letting young students use a spontaneous arithmetic language to understand the formal algebra, and it also falls in line with Dörfler (1991), who advocates a mutual relationship between theoretical and empirical generalizations.

This teacher-focused classroom design research had led to certain insights. This research is complementing traditional educational design research, and the chosen methodology shows the importance of having teachers as participating actors throughout the entire process. The teachers were met with surprising challenges in their teaching; for example, they challenged the students to justify what is behind a general formula. These challenges had consequences for the ongoing discussions in the design process, and as a result, the participating teachers changed how they talk about generalization. This shows that teachers as participating actors in all phases in educational design research can contribute to what are regarded as missing links in some design researches, both in terms of interweaving research findings and practice (Marotti et al. 2018) and in terms of the focus on instructional practice of the teacher (Cobb et. al. 2017).

Another reflection is that the teachers showed greater resistance to working with functional thinking than I had expected. However, it is important to note that, at the end of the process, the teachers changed their mind and talked about functional thinking as a new representation or a new tool for discussing generalizations in patterns with students. The teachers described their learning and the students learning about generalizations in relation to patterns as an "aha! experience" – patterns and the linear equation made for a new and an unexpected combination.

The challenges and the resistance in the process created opportunities, thus leading to the argument that challenges may be necessary for developing teaching.

Although the DPs are available, the results indicate that the teachers need to be challenged and negotiate the meaning of the DPs. This could be the consequence of the teachers having certain doubts in relation to the current DPs, or it may be an indicator of the more general importance of being challenged in order to develop new practices. The result supports that some forms of boundary objects (Wenger, 1998) appear necessary for a design research process. The analytical frame (Wenger, 1998) and the theoretical frame for the intervention – the DPs – made it possible to shed light on these three teachers' learning about teaching patterns with a focus on generalizations.

References

Beckmann, S. & Izák, A. (2015). Two perspectives on proportional relationships: Extending complementary origins of multiplication in terms of quantities. *Journal for Research in Mathematics Education 46,* 17–38.

Beckmann, S. (2018). *Mathematics for elementary teachers* (5th ed.). Boston, MA: Pearson.

Blanton, M. & Kaput, J. (2004). Elementary grades students' capacity for functional thinking. *Proceedings of the 28th Conference of the International Group of Psychology of Mathematics Education, 2*, 135–142.

Blanton, M., Brizuela, B. Gardiner, A. Sawrey, K. & Newman-Owens, A. (2015). A learning trajectory in 6-years-olds' thinking about generalizing functional relationships. *Journal for Research in Mathematics Education, 46*(5), 511–558.

Blanton, M., Levi, L., Crites, T., & Dougherty, B. (2011). Developing essential understanding of algebraic thinking for teaching mathematics in Grades 3–5. In R. M. Zbiek (Ed.), *Essential understanding series*. Reston, VA: National Council of Teachers of Mathematics.

Blanton, M., Stephens, A. Knuth, E., Gardiner, A. M., Isler, I., & Kim, J. S. (2015).. (2015). The Development of Children's Algebraic Thinking: The Impact of a Comprehensive Early Algebra Intervention in Third Grade. *Journal for Research in Mathematics Education*, *46*, 39–87.

Carraher, D. W., & Schliemann, A. D. (2015). Powerful ideas in elementary school mathematics. In L. D. English & D. Kirshner (Eds.). *Handbook of international research in mathematics education* (3rd ed., 191–218). New York: Taylor & Francis.

Carraher, D. W., Schliemann, A. D. & Schwartz, J. L. (2008). Early algebra is not the same as algebra early. In J. Kaput, D. Carraher, & M. Blaton, (Eds.), *Algebra in the early grades* (235–272). Mahwah, NJ: Lawrence Erlbaum/Taylor & Francis Group; Reston, VA: National Council of Teachers of Mathematics.

Caspi, S. & Sfard, A. (2012). Spontaneous meta-arithmetic as a first step toward school algebra. *International Journal of Educational Research*, 45–65.

Cobb, P., Jackson, K. & Dunlap Sharpe, C. (2017). Conducting Design Research Studies to Investigate and Support Mathematics Students' and Teachers' Learning. Research on the Teaching and Learning of Proof: Taking Stock and Moving Forward. In J. Cai (Ed.) *Compendium for research in mathematics education*. National Council of Teachers of Mathematics, 208–233.

Cobb, P. & McClain, K. (2001). An approach for supporting teachers' learning in social context. In F. L. Lin & T. Cooney (Eds.), *Making sense of mathematics teacher education* (207–232). Dordrecht, The Netherlands: Kluwer Academi.

Dörfler, W. (1991). Forms and means of generalization in mathematics, in A. J. Bishop (ed.), *Mathematical Knowledge: Its Growth through Teaching,* Kluwer Academic Publishers, Dodrecht, 63–85.

Floyd, A. & Arthur, L. (2012). Researching from within: external and internal ethical engagement. *International Journal of Research & Method in Education, 35*(2), 171–180.

Goodchild, S. (2011). Classroom research: impact and long term effect versus justice, liberation and empowerment? In L.O Björkvist & A. Röj-Lindberg (Eds.). *Long-Term Research in the Didactics of Mathematics and Science.* (9–25). Åbo: Faculty of Education. Åbo Akademi University.

Gravemeijer, K. & van Eerde, D. (2009). Design Research as a Means for Building a Knowledge Base for Teachers and Teaching in Mathematics Education. *The Elementary School Journal,* 5, 510–524.

Greeno, J. G. (2006). Theoretical and Practical Advances through Research on Learning. In J. L. Green, G. Camilli & P. B. Elmore (Eds.) *Handbook of Complementary Methods in Education Research* (795–822). Washington, DC: American Educational Research Association.

Johansson, M. (2006). *Teaching Mathematics with Textbooks a Classroom and Curricular Perspective*. Doctoral dissertation. Luleå: Luleå University of Technology Department of Mathematics.

Kaput, J. J. (2008). What is algebra? What is algebraic reasoning? In J. J. Kaput; D. W: Carraher, & M. Blanton (Eds.), *Algebra in the early grades* (5–17): Mahwah, NY: Lawrence Erlbaum/Taylor & Francis Group; Reston, VA: National Council of Teachers of Mathematics.

Kieran, C. (2004). Algebraic Thinking in the Early Grades: What Is It? *The Mathematics Educator*. 139–151.

Konrad, U. & Bakker, A. (2018). From implementer to co-designer: A teacher's changing role in a design research project. In A. Bakker (Ed.) *Design Research in Education A Practical Guide for Early Career Researchers* (246–254). Oxon, UK: Routledge.

Mariotti, M. A., Durand-Guerrier, V., Stylianides, G. J. (2018). Argumentation and proof. In T. Dreyfus, M. Artigue, D. Potari, S. Prediger & K. Ruthven (Eds.), *Developing research in mathematics education. Twenty years of communication, cooperation and collaboration in Europe* (75–89). Oxon, UK: Routledge.

McKenny, S. & Reeves, T. (2012). *Conducting Educational Design Research*. London: Routledge.

Mulligan, J., Cavanagh, M., Keanan-Brown, D. (2012). In B. Atweh, M. Goos, R. Jorgensen & D. Siemon, (Eds.). *Engaging the Australian National Curriculum: Mathematics – Perspectives from the Field*. Mathematics Education Research Group of Australian (47–70).

Mulligan, J., Mitchelmore, M. C., English, L. D. & Crevensten, N. (2013). Reconceptualizing Early Mathematics Learning: The Fundamental Role of Pattern and Structure. In L. D. English & J. T. Mulligan (Eds.) *Reconceptualizing Early Mathematics Learning, Advances in Mathematics Education,* New York: Springer.

NCTM (2008). *Navigating through reasoning and proof in Grades 9–12*. National Council of Teachers of Mathematics Reston, VA: NCTM.

National Agency for Education (2017). *Curriculum for the primary school, preschool class and leisure time centre* 2017. Stockholm: National Agency for Education.

Schliemann, A. D., Carraher, D., & Brizuela, B. M. (2007). *Bringing out the algebraic character of arithmetic: From children's ideas to classroom practice*: Mahwah: Lawrence Erlbaum and Associates.

Stephan, M. L. (2015). Conducting classroom design with teachers. *ZDM - Mathematics Education, 47*, 905–917.

Stephan, M., & Akyuz, D. (2012). A proposed instructional theory for integer addition and subtraction. *Journal for Research in Mathematics Education, 43*(4), 428–464.

Sterner, H. (2015). *Problematisera "görandet", Lärares lärande om kommunikation och resonemang i matematikundervisningen i en organiserad praktikgemenskap*. Licentiat Dissertation. Växjö: Linnaeus University.

Stylianides, G. J. & Silver, E. A. (2009). Reasoning-and-Proving in School Mathematics, The Case of Pattern Identification. In D. A. Stylianou, M. L. Blanton & E. J. Knuth (eds.), *Teaching and learning proof across the grades: A K–16 perspective*. New York, NY: Routledge.

Swedish Research Council (2017). *God Forskningssed*. Stockholm: Vetenskapsrådet.

Sztajn, P., Borko, H. & Smith, T. M. (2017). Research on Mathematics Professional Development. In J. Cai (Ed.) *Compendium for research in mathematics education*. National Council of Teachers of Mathematics (793–823).

Wenger, E. (1998). *Communities of Practice: Learning, Meaning, and Identity* (Learning in Doing: Social, Cognitive and Computational Perspectives). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511803932

Design and evaluation of practice-oriented materials fostering students' development of problem-solving competence: The case of working backward strategy

Ana Kuzle

University of Potsdam, Germany

In a design-research project on problem-solving, theory-based and practice-oriented materials were developed with the goal of fostering systematical development of students' problem-solving competence in a targeted manner by learning heuristics. Special attention was given to working backward strategy, which has been shown difficult for students to learn and use. In the study, 14 Grade 5 students participated in explicit heuristic training. The results show that even though the students intuitively reversed their thought processes before the explicit training, they experienced difficulties when solving complex reversing tasks, which improved considerably after explicit heuristic training. Thus, the study results showed that the developed materials using design-based research-approach promoted the development of students' flexibility of thought when problem-solving by working backward. At the end of the paper, the results are discussed with regard to their theoretical and practical implications.

Keywords
design research,
mathematics education,
problem-solving,
working backward strategy,
reversibility,
explicit heuristic training

Correspondence
kuzle@uni-potsdam.de

DOI
https://doi.org/10.31129/LUMAT.7.3.401

1 Introduction

Problem-solving is a binding process standard in different educational systems (e.g., Finnish National Board of Education [FNBE], 2004, 2014; National Council of Teachers of Mathematics [NCTM], 2000; The Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany [KMK], 2005) that is often neglected in school mathematics (e.g., Gebel & Kuzle, 2019). The plethora of research on problem-solving undergoing since the 1970s identified pivotal practices for problem-solving instruction (e.g., Grouws, 2003; Kilpatrick, 1985; Lester, 1985). Despite more than five decades of this accumulated knowledge, both empirical studies, as well as large-scale studies (e.g., PISA, TIMS study), reported that students are often unable to solve problem tasks. Moreover, teachers lack practical teaching materials to foster students' development of problem-solving competence and at the same time to consolidate their competence in the area (Gebel, 2015; Gebel & Kuzle, 2019; Kuzle & Gebel, 2016). In the context of this reform agenda, collaborative work between educational researchers and practitioners working on issues of everyday practice is crucial in order to overcome the gap between theory and practice (Jahn, 2014). Design-based research (DBR) as a research paradigm that brings the two poles – theory and practice – together, may help overcome this gap (Wang & Hannafin, 2005). In order to close the above-described problem-solving gap, theory-based and practice-oriented materials for middle-grade students were developed in the SymPa project (Systematical and material-based development of problem-solving abilities) in accordance with DBR methodology (Gebel, 2015; Kuzle, 2017a, 2017b; Kuzle & Gebel, 2016). The goal of the SymPa project was to promote the development of problem-solving abilities in Grades 4-6 (Gebel, 2015; Kuzle, 2017a, 2017b; Kuzle & Gebel, 2016).

In design research, depending on the existing object of investigation and the associated restrictions, different design aspects can be focused on that consequently allow to understand possible connections with regard to the fulfilment of the function of the design in a multifaceted way (Jahn, 2014). According to Collins, Joseph, and Bielaczyc (2004), different aspects are relevant for the multi-perspective educational design analysis: cognitive level, interpersonal level, group or classroom level, resource level, and institutional or school level. During the first seven DBR cycles within the SymPa project, the project evaluation focused on developing suitable and sustainable problem-solving materials for their implementation in practice (Kuzle, 2017b; Kuzle & Gebel, 2016) as well as on identifying the design elements contributing to the improvement of the problem-solving competence (Kuzle, 2017a). In other words, the resource level of the educational design was in the foreground of the analysis (Collins et al., 2004). At the same time, the project was analyzed with respect to factors, and conditions that favored and hindered the implementation of the materials in practice on the basis of two DBR cycles (Kuzle, 2017b). Thus, the institutional level of the educational design was analyzed (Collins et al., 2004). Hence, the first phase of the project had more practical output within educational design research.

Relating to the motive of enhancing the quality of research findings, the focus of this paper lies on another aspect relevant to educational design research, namely on the cognitive level (Collins et al., 2004). Specifically, the question about how the design of theory-based and practice-oriented materials for systematical development of mathematical problem-solving competence as well as on how explicit heuristic training organized around these materials affect the thinking and learning of participants over time, and subsequently their increase of knowledge in the context of mathematical problem-solving. This is exemplarily shown with respect to the strategy of working backward, which has been shown difficult for students to learn and use (Aßmus, 2010a, 2010b), albeit its potential in mathematics lessons and importance in everyday life. Through reversible thinking, an individual is capable of seeing things not only from one single perspective but also its reversal. It may also minimize both errors in every decision as well as the error of answers as students tend to review their answers by reversing the result to the initial value of the problem. Lastly, thinking reversible is one of the primary requirements to solve mathematical problems (Bruder & Collet, 2011; Krutetskii, 1976; Lompscher, 1975).

In the following sections, I outline relevant theoretical and methodological underpinnings for systematical development of problem-solving competence in the context of working backward, before showing how these got integrated into students' problem-solving material. On the basis of the educational design research, the development of students' ability to use the strategy of working backward when problem solving is presented. In the last section, I discuss the findings with regard to their theoretical and practical implications.

2 Theoretical foundation

2.1 Mental agility

Schoenfeld (1985) defined the concept of a problem as a subjective assessment:

> The difficulty with defining the term problem is that problem solving is relative. The same tasks that call for significant effort from some students may well be routine exercises for others, and answering them may just be a matter of recall for a given mathematician. Thus, being a 'problem' is not a property inherent in a mathematical task. Rather, it is a particular relationship between the individual and the task that makes the task a problem for that person. (p. 75)

Similarly, Bruder and Collet (2011) define the concept of the problem as person-dependent. For them, a task becomes a problem for an individual when it seems unfamiliar, and when a promising solution is not immediately at hand. In that manner, problem-solving refers to a directed cognitive process in which the problem solver determines how to overcome an individual barrier resulting from bringing the initial state to the target state (Bruder & Collet, 2011; Schunk, 2008).

Problem-solving competence relates to cognitive (here heuristic), motivational and volitional knowledge, skills and actions of an individual required for independent and effective dealing with mathematical problems (Bruder & Collet, 2011). Accordingly, each individual must develop the ability to solve problems independently (e.g., Gebel & Kuzle, 2019; Stanic & Kilpatrick, 1989), and should learn approaches (heuristics) for solving mathematical problems and how to apply them in a given situation, develop reflectivity on own actions, and develop willingness to work hard (KMK, 2005; NCTM, 2000).

Research (e.g., Carlson & Bloom, 2005; Schoenfeld, 1985) showed that problem-solving activities in mathematics require skills and understanding that are often not readily apparent to novice problem solvers compared to experienced problem solvers. Especially, intuitive problem solvers possess particular mental agility (Liljedahl, Santos-Trigo, Malaspina, & Bruder, 2016), which is fundamental to mathematical problem-solving. Lompscher (1975) defined the concept of *mental agility* as a performance characteristic of the individual, which provides both quantitative as well as qualitative characteristics, which influence the intellectual activity. By this, Lompscher (1975) understands the ability to analyze the objective reality of the subject. Consequently, mental agility develops from this very activity which the subject exercises in the mental process in interrelation with objective reality. Accordingly, the mental mastery of the performance requirements of objective reality is expressed through the abilities of the subject. Lompscher (1975) described the mental agility through three subdomains. First of all, the mental operation is considered, which contains solidified action sequences. Processed knowledge forms networks in the long-term memory and, when applied, characterizes its quality (course quality). The aforementioned forms the second subdomain of mental agility and is supplemented by the willingness to actively apply one's knowledge (attitude) (Lompscher, 1975). Here, I mainly limit myself to the first two subdomains.

Mental operations concretize every cognitive activity. Regardless of the object of the action – the goal or the content of the action – they form complex sequences of operations. This results in mental operations, among other things, during the examination of things and characteristics as well as in problem-solving situations (Lompscher, 1975). In addition, every mental action is characterized by content (e.g., concepts, connections, procedures), process (e.g., systematic planning, independence, accuracy, agility), and partially conscious goals and motives. One of the most important mathematically relevant progression qualities is mental agility. According to Lompscher (1975), "flexibility of thought" expresses itself

> … by the capacity to change more or less easily from one aspect of viewing to another one or to embed one circumstance or component into different correlations, to understand the relativity of circumstances and statements. It allows to reverse relations, to more or less easily or quickly attune to new conditions of mental activity or to simultaneously mind several objects or aspects of a given activity. (p. 36)

The flexibility of thought is expressed by one' ability to:

1. reduce a problem to its essentials or to visualize it by using visual and structuring aids, such as informative figures, tables, solution graphs or equations (*reduction*),
2. reverse trains of thought or reproduce these in reverse, such as by working backward (*reversibility*),
3. simultaneously mind several aspects of a given problem or to easily recognize any dependencies and vary them in a targeted manner (e.g., by composing and decomposing objects, by working systematically) (*minding of aspects*),
4. change assumptions, criteria or aspects in order to find a solution, such as by working forward and backward simultaneously or by analyzing different cases (*change of aspects*), and
5. transfer an acquired procedure into another context or into a very different one by using analogies, for instance (*transferring*).

These manifestations of mental agility can be related to *heurisms*, which are known from the analyses of Pólya's approaches (1945/1973). *Heuristics* can be defined "as kinds of information, available to students in making decisions during problem solving, that are aids to the generation of a solution, plausible in nature rather than prescriptive, seldom providing infallible guidance, and variable in results" (Wilson, Hernandez, & Hadaway, 1993, p. 63). Moreover, not only the knowledge of different heuristics (flexibility of thought) is needed when problem-solving, but also self-regulatory abilities which evolve gradually through a 5-phase model (Zimmerman, 2002). It has been a long-term goal of mathematics educators to provide students with the skills necessary for success in problem-solving.

2.2 Reversibility

Problem-solving by working backward describes the ability to reverse trains of thought or reproduce these in reverse (e.g., Liljedahl et al., 2016; Pólya, 1945/1973). Other than when working forward, the target state forms the starting point in the solution process, whereas the calculated value forms the initial value of the problem. Thus, working backward leads to the entire thought process being reversed, since the task no longer corresponds to working forward (changing the direction of processing) (Bruder & Collet, 2011). Aßmus (2010a, 2010b) distinguished between two aspects of reversibility. On the one hand, the operation as such can be reversed (e.g., '+' becomes '–'), and, on the other hand, the sequence of the task processing can be reversed.

The ability to independently reverse trains of thought when working on mathematical problems has been recognized as one of the indicators for identifying mathematically gifted children (Käpnick, 1998). Consequently, the research on reversibility during problem solving has been primarily done with gifted and talented students. For instance, while working with potentially talented primary grade students, Aßmus (2010a, 2010b) investigated performance in heterogeneous primary mathematics lessons (Grade 2 and Grade 4) when solving reversing tasks. The results showed that the basic understanding of reversal was not present from the beginning. While many gifted Grade 2 students still had problems, on average, they were much more successful than the average children of the same age. For instance, no student from a control group was able to solve a problem with an unknown initial state, whereas at least 9% of the potentially gifted Grade 2 students (N = 182) succeeded in completing the tasks correctly, and in 35% of cases reasonable solution approaches could be identified. Symbolic tasks were generally processed backward more intuitively than word problems in which starting from an unknown initial state, various transformations needed to be performed in order to determine the initial state. Difficulty in the processing of the latter was maintaining the correct reversal of the operations or taking all operations into account when reversing.

On the other hand, Grade 4 students were more successful, with 36% of students reaching the correct solution (Aßmus, 2010b). Thus, reversibility was differently pronounced by primary grade students. Aßmus (2010b) hypothesized that this ability develops in the course of years with respect to average students (Aßmus, 2010b), though it may be more characteristic for gifted students (Aßmus, 2010a, 2010b).

The latter finding was supported by Amit and Portnov-Neeman (2016) who examined the effect of explicit teaching of problem-solving strategies, with a special focus being given to working backward strategy, on the ability of mathematically talented Grade 6 students to recognize and solve reversing tasks. The group that received explicit training showed higher results than the control group. Here, the students from the experimental group showed a better, clearer understanding of the strategy, and the strategic use improved over time (Amit & Portnov-Neeman, 2016). Moreover, they were more much resourceful in their solutions when solving a wide variety of reversing tasks. Amit and Portnov-Neeman (2016), therefore, confirmed the results of Aßmus (2010a, 2010b).

Whereas Aßmus (2010a, 2010b), and Amit and Portnov-Neeman (2016) focused on reversible thinking ability of gifted and talented students in the context of mathematics problem solving, Gullasch (1967) examined the relationship between mathematical problems and mathematical ability of Grade 7 students. His study revealed a high correlation between the reversibility of the mental activity, the level of school performance, and the ability to abstraction. Accordingly, the reversal of solution paths sets a basic level of mental activity.

2.3 Learning problem solving

In the field of problem-solving, there are two different approaches to learning heuristics. In an implicit heuristic training, it is assumed that the students internalize and unconsciously apply strategies they have learned through imitating practices of the teacher, and through sufficient practice. On the other hand, explicit heuristic training refers to making a given heuristic a learning goal, which is practiced step by step (e.g., Schoenfeld, 1985). Bruder and Collet (2011) pursued an explicit heuristic training focusing around Lompscher's (1975) idea of "flexibility of thought" in combination with self-regulation (Zimmerman, 2002), which consisted of the following five phases:

1. *Intuitive familiarization*: The teacher serves as a role model when introducing a problem to the students. This is achieved through moderation of behaviors by engaging in self-questioning (e.g., "What is the problem asking for?" "What information am I given?" "Am I headed in the right direction?") pertaining to different phases of the problem-solving process (before, during, and after) (Kuzle & Bruder, 2016). At this point, the heurism in focus is not specified.
2. *Explicit strategy acquisition:* The students get explicitly introduced to the heurism in focus by reflecting on the first phase, namely the particularities of the heurism get discussed, and the heurism is given a name. Here a prototypical problem for the heurism in focus is used so that the students can more easily recognize and use the heurism in future tasks.

3. *Productive practice phase:* The students practice the newly acquired heurism by solving different problems. These do not reproduce type problems, but rather expand the possibilities from the first two phases. Differentiation is a guiding concept so that students can choose at what cognitive level they want to work and adapt the observed learning behavior.k
4. *Context expansion*: The students practice the use of heurism in focus independent of a mathematical context. In that way, the students learn to flexibly, intuitively and independently of a context use the heurism in focus.
6. *Awareness of own problem-solving model:* The students reflect on their problem-solving process and document it.

Untrained problem solvers are often unable to consciously access the above-outlined flexibility qualities (Bruder & Collet, 2011; Liljedahl et al., 2016). In their research at the lower secondary level, Bruder and Collet (2011) were able to show that less flexible students (e.g., students with difficulties in reversing thought processes or transferring an acquired procedure into another context) profit from explicit heuristic training. Concretely, they were able to solve the problems just as well as more flexible students, who solved the problems intuitively. Thus, the problem-solving ability can be acquired through the promotion of manifestations of mental agility (reduction, reversibility, minding of aspects and change of aspects) in combination with self-regulation.

2.4 Design-based research in the context of SymPa project

Learning is a complex process, which depends on many factors, and thus, is difficult to control. Design-based research (DBR) as a research paradigm offers the opportunity to develop innovative teaching practices, and to develop context-sensitive learning environments. According to Wang and Hannafin (2005), design-based research is "a systematic but flexible methodology aimed to improve educational practices through iterative analysis, design, development, and implementation, based on collaboration among researcher and practitioners in real-world settings, and leading to contextually-sensitive design principles and theories" (p. 6-7). Hereby, they especially underline the flexible character of DBR and the importance of synergy of theory and practice, in contrast to other research paradigms.

The Design-Based Research Collective ([DBRC], 2003) lays down the cyclical and continuous nature of DBR comprising of design, enactment, analysis and re-design phase (see Figure 1). Under design is theory-driven development of a teaching-learning environment or a material understood, which will be implemented in the phase of enactment. In the next step, designed teaching-learning environment or material is analyzed in the evaluation phase. The improvements get then implemented in a re-design phase, and the cycle starts from the beginning on. In that manner, the result of any DBR approach is the development of new knowledge or suggestions on how to improve educational practice(s), such as exploring possibilities for creating novel learning and teaching environments, developing contextual theories of learning and instruction, advancing and consolidating design knowledge, and increasing the capacity for educational innovation (e.g., Collins et al., 2004; DBRC, 2003).

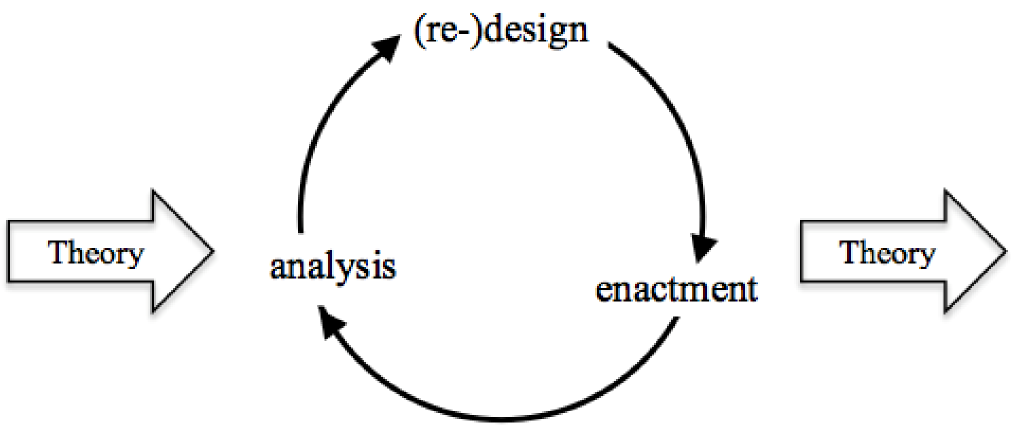


Figure 1. DBR cycle.

During the first seven DBR cycles of the SymPa project, the evaluation focused on analyzing to what extent are theory-based and practice-oriented problem-solving materials suitable and sustainable for their implementation in practice (Kuzle & Gebel, 2016) as well as on identifying the design elements contributing to improvement of the problem-solving competence (Kuzle, 2017a). Kuzle and Gebel (2016) reported that it was possible to develop a curriculum that met the local demands which enabled the implementation of problem-solving in practice. As a result, context-related design principles for the development of problem-solving material for Grade 6 students were developed (Kuzle, 2017a). The results showed that students needed an emotional incentive (hereby the figures) in order to be willing to solve problems and to prompt their reflective behaviors. Transparency of the material structure supported students' independent work, whereas material design (differentiation, transparent material structure with explicit reflections) was an important factor in the development of self-regulatory processes when problem-solving. Lastly, various design elements (text information, sample problem) allowed for explicit strategy acquisition and 3 to 4 problems seemed optimal for flexibility use. Moreover, Kuzle and Gebel (2016) demonstrated that DBR paradigm allowed creating novel teaching environments in which theory and practice were not detached from one another, but rather complemented each other (resource level of the educational design) (Collins et al., 2004). At the same time, the design was analyzed with respect to different objectives (e.g., language, level of performance, learning pedagogies) and subjective factors (e.g., school and personal influences) which inhibited full-implementation of the curriculum (institutional level of the educational design) (Collins et al., 2004). However, how these materials affect students' thinking and learning over time and subsequently their increase of knowledge in the context of mathematical problem solving remained open (cognitive level of the educational design) (Collins et al., 2004).

3 Research questions

On the basis of the above theoretical considerations and empirical results, the following research questions guided the study on problem-solving by working backward:

1. How do Grade 5 students solve reversing tasks before and after explicit heuristic training?
2. To what extent are Grade 5 students able to solve reversing tasks by working backward before and after explicit heuristic training?

4 Method

4.1 Research design and sample

For this study, an explorative qualitative research design was chosen. The study participants were Grade 5 students who showed interest in attending additional mathematics lessons on a voluntary basis that focused on problem-solving. In total 14 students (nine girls and five boys) from one rural school in the federal state of Brandenburg (Germany) participated in the study, and thus attended explicit heuristic training on the working backward strategy. Their performance in regular mathematics classes was good to very good.

4.2 Context of the study

The WoBa study (Problem Solving by Working Backward) was embedded in the SymPa project during which the students participated in explicit heuristic training which lasted about 7 months (see Figure 2). The explicit heuristic training took place once per week (45-minute lesson). During this period, the students received explicit problem-solving instruction pertaining to different heuristic auxiliary tools (informative figure, table), heuristic strategies (working systematically, working forward, working backward, analogy), and heuristic principles (composing and decomposing, invariance), which lasted two to three lessons per heurism. The lessons were taught by an experienced mathematics teacher.

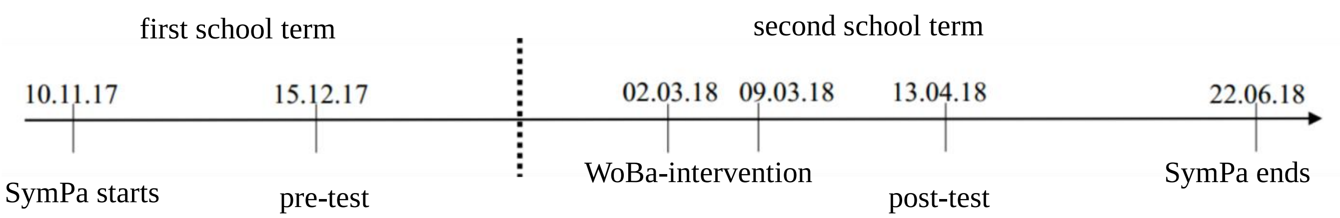


Figure 2. Timeline of the SymPa project.

During the explicit heuristic training, the students systematically learned heurisms using theory-based and practice-oriented materials (Kuzle, Gebel, & Conradi, 2017-2019) on the basis of the problem-solving teaching concept of Bruder and Collet (2011). The problem-solving material focusing on the working backward strategy is outlined below.

In the phase of intuitive familiarization, the students are given a representative problem for the working backward strategy (see Figure 3), which is solved together with the teacher, who serves as a moderator. Here the imitation of teachers' behavior takes place through self-questioning.

3.2 Working backward

3.2.1 Misplaced glasses

Oh Profi, where are your glasses?

I don't know. I must have misplaced them.
I've been thinking the entire time about what I've done today.

a) What does Profi mean by that?
b) How can he find his glasses again?

Figure 3. Introduction task on the working backward strategy (Kuzle et al., 2017-2019).

In the phase of explicit strategy acquisition, the working backward strategy gets formally introduced through a short student-centered information text (see Figure 4), and a partially worked out example (see Figure 5).

What is working backward?
Working backward is closely related to working forward, but runs in the other direction.
Here we start from the target state and follow the path to the initial state.
Questions, such as "What is wanted?", "What do I know about what I am looking for? "What do I need in order to find what I'm looking for?" offer orientation.

Figure 4. Information text on the working backward strategy (Kuzle et al., 2017-2019).

Example
Probi found this riddle in a magazine:

With a number between 1 and 9, six arithmetic tasks, starting with the upper result, are to be solved one after another in a clockwise direction in order to arrive at the final results of 136.

I'll try out some numbers.

That would take a really long time.
What strategy did we (just) learn here?
I would start with 136 and solve the task the other way round.

Figure 5. An example illustrates the working backward strategy (Kuzle et al., 2017-2019).

In what follows, at least three reversing tasks of different cognitive levels are presented (productive practice phase). This allows differentiation, where each student solves as many problems as he or she can. In addition, problems from different mathematical content areas are covered to allow transfer (context expansion phase).

Lastly, the tasks stipulate students to reflect on their problem-solving process (see Figure 6).

3.2.2 5-Aunts

Probi always gets candy from his aunts when he visits them. Each aunt gives him as much candy as he already has and one more. Probi has 5 aunts. After visiting all of them has 127 candies in his bag.

a) How many candies did he have before visiting them?
b) Have you been able to work backward on this task? Why did this strategy fit the task?
c) What heuristic tools did you use?

3.2.3 Number crusher

Input → Number crusher → Output

The "Number crusher machine" processes the numbers 1, 2, 3, and so on.
Even numbers are halved, uneven numbers are reduced by 1, e.g., 6 → 3 und 5 → 4.
The output number is then put back into the „Input" until 0 becomes the "Output", e.g., 5 → 4 → 2 → 1 → 0. So for 5 you would beed four steps (→) to reach 0.
Therefore, 5 is calles a 4-step number.

a) Examine how many steps you need for other numbers to reach 0!

For this you can use one of the learned heuristic tools.

b) How many 4-step numbers are there? List all of them.

Here by using a table, as a heuristic tool, can discover some interesting things.

3.2.4 Cutting paper

Profi and Probi play a game: Profi fold a piece of rectangular paper and then makes a straight cut. Probi only sees the end product and should find out how Profi folded and cut the paper.

a) Find out how Profi folded and cut the paper.
b) Did you work backward in part a)? Why?
c) What other heuristic strategies would fit you approach in a)?

3.2.5 Pouring water

Probi suggest Profi a bet and gives him two buckets:
„This is a 3-liter and a 5-liter bucket. They don't have any markings. You can now pour as much water back and forth as you want until you have exactly 4 liters of water in the 5-liter bucket. I bet you a hot chocolate."

a) Who will get the hot chocolate? Why?

Try to be systematic!
Think before you start solving the problem.

b) What heuristic tool did you use?
c) Try the informative figure and the table.
Which heuristic tool did you find best here? Why?

Figure 6. Further tasks with respect to the working backward strategy (Kuzle et al., 2017-2019).

From a design perspective, two figures, namely Profi (shape of an exclamation mark) and Probi (shape of a question), are introduced to support students' willingness to work hard. While Probi asks questions and gets stuck like every novice problem solver, Profi represents an expert problem solver who offers support to novice problem solvers (i.e., students).

4.3 Data collection instruments and procedure

The study data consisted of (1) written data, (2) oral data, and (3) observations. (1) The written data were comprised of the students' solutions to pre- and post-test, each containing three problems addressing different mathematical levels. The pre-test was issued in December before that explicit training on working backward strategy at the beginning of the year started, whereas post-test one month after the end of the explicit training (see Figure 2). The students were allotted 45-minutes for each test. Both tests contained partially similar tasks in order to be able to compare the students' development of the working backward strategy use (see Table 1). The pre- and post-test included two types of reversing tasks. On the one side, the tests included tasks with the unknown initial state. In this case, starting from an unknown initial state, various transformations and the final state are described, and the initial state needed to be determined. The complexity of tasks varied based on the number of different operations, namely one ('Candy task') and two ('Four gates' task', 'Devil's task') operations. On the other hand, the tests included tasks that required a flexible reversal of relations. Thus, tasks which cannot be solved by working exclusively by working forward or backward, but whose processing requires flexible handling of relations which are often reversed several times, provided that they are not only tried out ('Circle task', 'Dogbone task', 'Rectangle task'). Additionally, the students reflected on different problem-solving strategies using reflection sheets at the end of explicit training.

(2) For the purpose of gaining a detailed insight into students' problem-solving processes, a brief interview (5-minutes) was conducted with four individual students, who were chosen on the basis of their results on the pre- and post-test. The following questions served as guidelines: "How did you come up with the solution?" "Do you think you could have solved solve the problem in another way?"

(3) During the explicit training on the working backward strategy, the researcher observed the lessons and made observation notes. Multiple data sources were used to assess the consistency of the results, and to increase the validity of the instruments.

Table 1. Pre- and post-test tasks.

<table>
<tr><th>Pre-test tasks</th><th>Post-test tasks</th></tr>
<tr><td>

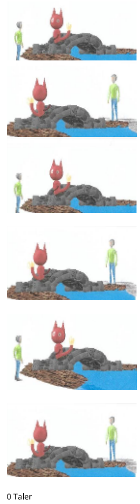
Four gates' task

A man goes apple picking. To take his harvest to the town, he has to pass through four gates. At each gate, there is a guard and demands half of his apples and one more apple. In the end, the only thing left to the man is an apple. How many apples did he have at the beginning?

</td><td>

Devil's task

The devil says to a poor man: "Every time you cross this bridge, I will double your money. But every time you come back, you have to throw eight thalers in the water." When the man returned for the third time, he did not have a single thaler left. How many thalers did he have at the beginning?



</td></tr>
<tr><td>

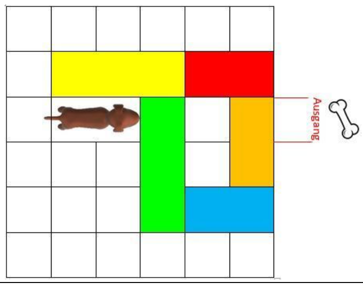
Candy task

Marie gets a bag of sweets from her grandmother as a present. On the first day, she eats half of the sweets. On the second day, she eats half of the remaining sweets. Afterwards, she only had six sweets left. How many sweets were in the bag at the beginning?

</td><td>

Dogbone task

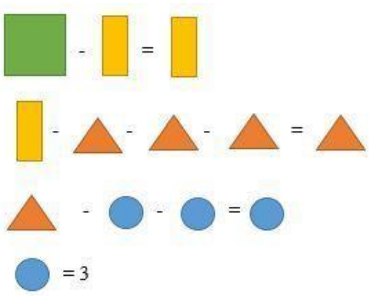
The dog wants to get to his bone. Unfortunately, the way is blocked by colored blocks. Can you bring the dog to his bone? Find a way.



</td></tr>
<tr><td>

Circle task

What are the numbers for the remaining pieces? Solve the calculation.



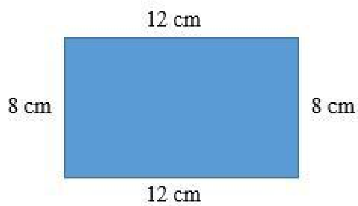
</td><td>

Rectangle task

The sides of the blue rectangle are a total of 40 cm long. The blue rectangle is to become two rectangles. The sides of the two rectangles should be 40 cm long in total.

a. What side lengths can the two rectangles have?

b. Can you find any other solutions?



</td></tr>
</table>

4.4 Data analysis

The analysis of the data was carried out in several steps. First, the data were examined with respect to used heurisms (1st research question). There was the possibility to solve the tasks with the help of working forward strategy (WF) or working backward strategy (WB). These could also have been processed without any specific strategy or not solved at all (NS). One coding was done per task, so that in the end, an overview was created, which reflected used problem-solving strategies of the respective student. The tasks were considered to be solved backward when the final result was recognized as the initial value, and the arithmetic operations of the task were correctly reversed ('Four gates task', 'Devil's task', 'Candy task', 'Circle task'). The 'Rectangle task' was also considered to have been solved backward, if it could be seen that the perimeter of the rectangle was divided by two and the result was distributed over the perimeter equation of a rectangle. The solution to the 'Dogbone task' was accepted as backward as long as it became clear that the block closest to the dog was moved. Subsequently, the inductive analysis of the problem-solving process was carried out by taking into account the different application performances of working backward strategy (2nd research question). In order to classify each student's achievement, these were assigned to individual levels of working backward (see Table 2). This was again carried out per task, in order to evaluate each student's progress in the project with respect to their ability to reverse trains of thought or reproduce these in reverse.

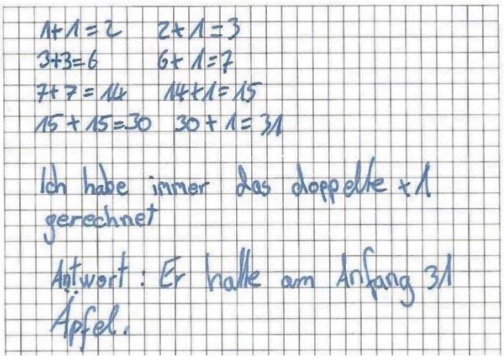
Table 2. A framework for different levels of working backward

Levels of working backward	Description of students' behavior
WB1	Students do not use the given target state as a starting value for the calculation. The required operations are not reversed.
WB2	Students are able to use the given target state as a starting value. The required operations are not reversed.
WB3	Students can correctly reverse the required operations. However, the task is not calculated to the end so that the correct result is not achieved.
WB4	Students are able to work backward correctly.

Some students' solutions to the 'Four gates task' with assigned levels of working backward can be seen in Figures 7 to 9.

"At the beginning, he had 16 apples."

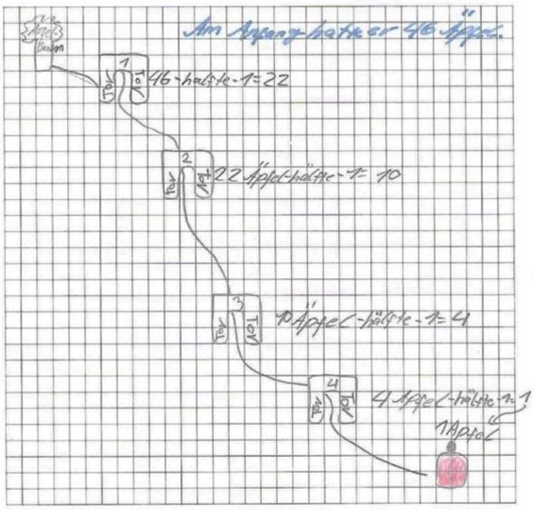
Figure 7. One student's solution performing on WB1 level.



"I always calculated the double +1."

Response: "He had 31 apples at the beginning."

Figure 8. One student's solution performing on WB2 level.



At the beginning he had 46 apples.
1. gate: "46 - half - 1 = 22"

2. gate: "22 apples - half - 1 = 10"

3. gate: "10 apples - half - 1 = 4"

4. gate: "4 apples - half - 1 = 1 apple"

Figure 9. One student's solution performing on WB4 level.

The students' self-reflection on different problem-solving strategies were also used to interpret the results as well as individual interviews. The latter was needed to correctly comprehend students' problem-solving processes.

5 Results

5.1 Students' strategies when solving reversing tasks

In Table 3, it can be seen that the majority of the students intuitively used the working backward strategy on almost all pre-test tasks. One student only (#12) consistently used the working backward strategy on all three pre-test tasks. Additionally, two students (#5, #9) used the working forward strategy when solving the 'Candy task' and the 'Circle task'. Two students (#7, #14) did not employ any strategies when solving the 'Four gates task' and the 'Circle task'.

Table 3. Classification of the students' solutions on the pre-test in relation to used heurisms (• 'Four gates task', • 'Candy task', • 'Circle task')

	Student													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
NS							•							••
WF					•				•			•••		
WB	•••	•••	•••	•••	••	•••	••	•••	••	•••	•••		•••	•

Similar results can be seen in Table 4 illustrating students' strategies on the post-test. Six students (#4-#9) employed the working forward strategy, whereas five of them when working on the 'Dogbone task'. Thereupon, it can be deduced that the reversal of thought processes has not proved to be a useful strategy for all students with respect to the 'Dogbone task'. Two students (#12, #13) did not employ any strategies when solving the 'Devil's task' and the 'Rectangle task'. Thus, after the explicit training, the majority of the students used the working backward strategy in most cases, and some were able to employ the most effective strategy for them.

Table 4. Classification of the students' solutions on the post-test in relation to used heurisms (• 'Devil's task', ○ 'Dogbone task', • 'Rectangle task')

	Student													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
NS												•	•	
WF				○	•	○	○	○	○					
WB	•○•	•○•	•○•	••	○	••	••	••	••	•○•	•○•	○•	•○	•○•

Though Table 3 and Table 4 illustrate students' ability to (intuitively) reverse their thought processes when solving reversing tasks, there were differences on students' level of problem-solving by working backward which are reported on in the next section.

5.2 Students' level of working backward when solving reversing tasks

Based on different application performances of the working backward strategy on both pre- and post-test, the students' solutions were sorted into the appropriate levels of working backward. In that manner, a rough overview of the differentiated performance of each student was created (see Table 5). The performance of the students varied greatly, however, the majority (N = 11) were able to solve the 'Candy task' correctly (WB4), which could be solved by a simple reverse operation. The 'Four gates task' has shown to be more difficult for students. In total three students (#2, #4, #11) were able to solve the problem correctly (WB4) (see Figure 9), whereas eight students performed on WB1 or WB2 level (see Figure 7 and Figure 8). It may be that such poor performance was due to its complexity, as the task required two combined inverse operations. Nine out of fourteen participants were also able to work intuitively on the 'Circle task' performing on WB4 level, whereas two students were on WB2 level. Only one student (#14) was not able to solve the task. This symbolic task demanded not only the reversal of the existing operations but also an extension of the part-whole relation.

Table 5. Overview of students' solutions on the pre-test in relation to levels of working backward (• 'Four gates task', • 'Candy task', • 'Circle task')

	Student													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
NS							•							••
WF					•				•			•••		
WB1			•		•				•				•	
WB2	•	•				••		•		•				
WB3			•											
WB4	••	••	•	•••	•	•	••	••	•	••	•••		••	•

The students' performance on the post-test was more homogeneous than on the pre-test (see Table 6). Each student solved at least one problem correctly using working backward strategy (WB4). The 'Dogbone task' was solved by six and three students performing on WB3 and WB4 level, respectively. The task demanded a high degree of mental agility as it was an open task, which can be solved with both working forward and backward strategy. The 'Rectangle task' was successfully solved by nine students, whereas three students (#1, #4, #6) only showed rudimentary approaches to working backward (WB1). Compared to other tasks, it demanded geometric knowledge, namely calculating the perimeter of a rectangle and taking into account geometric ratio distribution. Despite the complexity of the 'Devil's task' due to the two combined inverse operations, eight students were able to solve the problem correctly (WB4), and two performing on WB3 level. Still, one student (#12) was not able to solve this problem, and three students showed were rudimentary approaches (WB1 or WF).

Table 6. Overview of students' solutions on the post-test in relation to levels of working backward (● 'Devil's task', ○ 'Dogbone task', • 'Rectangle task')

	Student													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
NS												●	•	
WF				○	●	○	○	○	○					
WB1	● •			•		•				●				
WB2														
WB3			○						●	○	● ○	○	○	○
WB4	○	● ○ •	● •	●	○	●	● •	● •	•	•	•	•	●	● •

Lastly, Table 7 illustrates more closely the students' performance on similar tasks, namely the 'Four gates task' and the 'Devil task'. At the beginning of the explicit training, the performance of eight students corresponded to the two lowest levels of working backward. These students only partially used the given target state as their initial value, and there was no reversal of the operations (see Figure 7 and Figure 8). After the explicit training, the results of the post-test showed that 10 students were able to solve the task partially correct (WB3) or correct (WB4). All of these students reversed both operations in a proper manner. However, two students forgot a subtask (one more crossing of the bridge) and for that reason did not achieve the correct end result. Only two students (#2, #4) were able to solve both tasks correctly (WB4) before and after the explicit training. Overall, it can be said that the ability to problem solve by working backward increased in seven of the remaining 12 students. For instance, students #7 and #14 did not initially show any strategic approaches. At the end of the explicit training, they were able to complete the corresponding task by using the working backward strategy. A direct comparison between the other tasks on both tests was not possible, because they were neither similar in context nor structure.

Table 7. Comparison of students' performance on similar tasks (● 'Four gates task', • 'Devil's task')

	Student													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
NS							●					•		●
WF					•							●		
WB1	•		●		●				●	•			●	
WB2	●					●		●		●				
WB3									•		•			
WB4		●•	•	●•		•	•	•			●		•	•

The individual students' achievements are considered in Table 8 on the basis of a short interview. For this purpose, four children were selected, regardless of their gender. Additionally, Table 5 and Table 6 are compared and the self-assessment of the students is used as an interpretation aid.

Table 8. Assessment of the individual achievement when problem-solving by working backward

Student	Assessment of individual achievement with respect to the level of working backward
#2	In the pre-test, the student used the working backward strategy in all tasks, whereby she reached WB2 level once ('Circle task') and WB4 level twice ('Four gates task', 'Candy task'). In the post-test she solved all tasks by using the working backward strategy, performing on WB4 level. In the interview, it became clear that she already knew how to use the working backward strategy before the explicit training. For that reason, it was easy for her to solve the reversing tasks.
#8	In the pre-test, the students used the working backward strategy in all tasks, whereby he reached the WB2 level once ('Four gates task), and WB4 level twice ('Candy task', 'Circle task'). The results of the post-test showed performance improvement when solving similar tasks (from WB2 to WB4 level). In the interview, he reported that he found it difficult to work backward and was often confused when working on reversing tasks.
#12	In the pre-test, the students used the working forward strategy in all tasks. On the other hand, two tasks in the post-test were solved by using the working backward strategy, namely 'Dogbone task' (WB3) and 'Rectangle task'(WB4). Her solution to the 'Devil's task' did not allow any conclusions with respect to used problem-solving strategy. In the interview, she reported that she often got confused when working on the reversing tasks, and it was difficult for her to reproduce her thoughts in reverse.
#14	In the pre-test, the student worked on two tasks ('Four gates task', 'Circle task') without using a specific heuristic strategy, whereas she solved the 'Candy task' by using working backward strategy (WB4). In the post-test, she solved two tasks correctly by using the working backward strategy, namely 'Devil's task' and 'Rectangle task' (WB4). She also reversed her thoughts when working on the 'Dogbone task', but reached WB3 level only.

Reflection sheets of all students on the topic of "heuristic strategies" provided additional information on the extent to which they found the working backward strategy useful. Four students rated the strategy as easy, whereas seven students as difficult. The rest abstained from reflecting on the strategy. Moreover, six students reported that they would use this strategy in the future. Self-reflection of four selected students was consistent with their performance on the pre- and post-test.

6 Discussion and conclusions

For almost one school year, the students took part in explicit heuristic training. Within the WoBa study, a special focus was given to problem-solving by working backward, and the students' ability to work backward was evaluated. At the beginning of the explicit training, almost all students' approaches showed instances of working backward. They intuitively reversed their thought processes when solving the reversing tasks. Overall, every student used this strategy in the post-test, with as many tasks as in the pre-test were not solved backward. There may be various reasons for this. The students used both working forward and backward strategy when solving the logic task ('Dogbone task'). Consequently, no arithmetic or geometric skills were required, but only the reversibility of the trains of thought. On the other hand, this task can also have been perceived as difficult, since such a prototypical task was not dealt with within the explicit training (see Figures 3-6). Should the former be the case, then this would be a distinguishing feature with regard to mental agility and with it developed mathematical ability (Gullasch, 1967). Simple reversing task ('Candy task'), as well as the symbolic task ('Circle task'), were intuitively solved by using the working backward strategy compared to complex reversing tasks ('Four gates task', 'Devil's task').

Before the explicit training, the students were at different stages of development with respect to reversibility. Although the students found it difficult to reverse trains of thought and reproduce operations in reverse, the results show improvement with respect to different levels of working backward during problem-solving. After the explicit training, students were able to further develop their skills and reach the next stage of development (Bruder & Collet, 2011; Wygotski, 1964). It can be also assumed that the students were able to detach the structure of the heuristic strategy from the task context and transfer it to similar tasks (Lompscher, 1975). This became particularly evident when comparing students' performance on similar tasks, namely the 'Four gates task' and the 'Devil's task'. Consequently, it can be concluded that mathematically-interested students could train their mental agility, and thus, compensate for deficits in the area of mental activity by using the heurisms imparted by the SymPa project (Bruder, 2014; Kuzle & Bruder, 2016; Lompscher, 1975). Moreover, the ability to reverse thought processes or to reproduce these in reverse when confronted with reversing tasks may not only be reserved for gifted students (Aßmus, 2010a, 2010b) but also mathematically-interested students can successfully develop this ability. Also, the willingness to apply the heuristic strategies, which were mentioned in the students' self-reflection, helped promote the intellectual ability and compensate for deficits (Bruder, 2014; Lompscher, 1975).

In the SymPa project we have consciously decided to develop theory-based materials which were evaluated in a school setting in order to simultaneously (1) (develop) and evaluate the suitability of practice-oriented materials, (2) to develop a sustainable problem-solving teaching concept, and (3) to gain insights into individual students' learning processes. The latter relates to the cognitive level (Collins et al., 2004) of DBR. The results have shown that a synergy of the teaching concept and the problem-solving material allowed mathematically-interested students access to problem-solving, specifically to working backward strategy.

Despite positive results in the context of problem-solving by working backward, some drawbacks need to be discussed. This study was an exploratory qualitative study using a specific sample in the context of additional mathematics lessons on problem-solving on a voluntary basis. Hence, the results are limited to mathematically-interested Grade 5 students. Additionally, a small sample was used, so not all processes were reported. These limitations suggest a possible next step in research. Since the problem-solving materials were developed for Grade 5 and 6 students, future studies may look into the extent to which they are implementable in regular mathematics lessons rather than in special mathematics contexts. Since the tasks cover different mathematical areas, they may be implemented flexibly. Moreover, the effect of the materials on the development of all students' problem-solving competence, not only with respect to the strategy of working backward, is an area highly important to investigate taking into consideration mathematics standards worldwide (e.g., FNBE, 2004, 2014; KMK, 2005; NCTM, 2000). It is also questionable whether a long-term intervention on the subject of working backward would have influenced the students' results or promoted their mental agility in the area of logic tasks. Additionally, the pre- and post-test tasks were selected according to the mathematics curriculum (RLP, 2015). During the post-test, however, it became clear that not all students were proficient in calculating the perimeter of rectangles, as this was not the content of the previous school year. This aspect had a significant influence on the results of the post-test. Lastly, the pre- and post-test were deliberately structured in such a way that they did not only consist of similar tasks in order to avoid routine processing of the tasks. However, a deeper insight into the development of reversibility between pre-/post-testing would have been provided by other similar tasks. This direction may be fruitful for future studies.

The SymPa project demonstrated the fruitfulness of the synergy between practice (i.e., school, practitioners) and theory (i.e., research, university staff). From the perspective of practice, the school gained high-quality material on problem-solving which allowed supporting needs of students interested in mathematics. From the perspective of theory, a framework for different levels of working backward was developed. This may also be used by practitioners in order to evaluate students' levels of working backward as well as to promote their development of reversibility of thought. Additionally, the study findings reflect a great potential for problem-solving in school mathematics. Both the developed materials using design-based research-approach and the teaching of heurisms in the classroom stipulated the development of students' flexibility of thought when problem-solving by working backward. The majority of students improved their ability to work backward, progressing to the next or second next level. In addition, almost all students reached the highest level of working backward. Thus, the study results show that the theory-based and practice-oriented materials using DBR approach not only allow sustainable implementation in practice (Kuzle, 2017a, 2017b) but also promote the development of targeted problem-solving abilities. Further research, however, is needed to evaluate the utility of the materials with respect to general problem-solving ability.

That DBR as a research paradigm may support gradual improvements in both practice and theory, and that with it further theoretical and practical developments are possible, preclude no doubts.

References

Amit, M., & Portnov-Neeman, Y. (2016). 'Explicit teaching' as an effective method of acquiring problem solving strategies - the case of 'working backwards'. In T. Dooley & G. Guedudet (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education* (pp. 3792–3800). Dublin, Ireland: DCU Institute of Education and ERME.

Aßmus, D. (2010a). *Fähigkeiten im Umkehren von Gedankengängen bei mathematisch begabten Grundschulkindern* [Ability to reverse thought processes in mathematically gifted primary school children]. In A. M. Lindmeier & S. Ufer (Eds.), *Beiträge zum Mathematikunterricht 2010* (pp. 137–140). Münster, Germany: WTM-Verlag.

Aßmus, D. (2010b). Fähigkeiten im Umkehren von Gedankengängen bei potentiell mathematisch begabten Grundschulkindern [Ability to reverse thought processes in potentially mathematically gifted primary school children]. In T. Fritzlar & F. Heinrich (Eds.), *Kompetenzen mathematisch begabter Grundschulkinder erkunden und fördern* (pp. 45–61). Offenburg, Germany: Mildenberger Verlag.

Bruder, R. (2014). Fachdidaktisch und lerntheoretisch begründete Modelle zum Lehren und Erlernen von Heurismen im Mathematikunterricht [Subject and theory-based models for teaching and learning heurisms in mathematics lessons]. In F. Heinrich & S. Juskowiak (Eds.), *Mathematische Probleme lösen lernen. Ars Inveniendi et Dejudicandi* 3 (pp. 31–46). Münster, Germany: WTM-Verlag.

Bruder, R., & Collet, C. (2011). *Problemlösen lernen im Mathematikunterricht* [Learning problem solving in mathematics]. Berlin, Germany: Cornelsen.

Carlson, M. P., & Bloom, I. (2005). The cycle nature of problem solving: An emergent multidimensional problem-solving framework. *Educational Studies in Mathematics, 58*, 45–75.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design-based research: Clarifying the terms. Introduction to the learning sciences methodology strand. *The Journal of the Learning Sciences, 13*(1), 15–42.

Gebel, I. (2015). *Im Forderunterricht Problemlösen lernen und lehren: Design Based Research als Chance zur Entwicklung praxisbezogener und theoriegeleiteter Materialien* [Learning and teaching problem solving in demanding classes: Design Based Research as an opportunity to develop practice-oriented and theory-based materials] (Unpublished master's thesis). University of Osnabrück, Osnabrück, Germany.

Gebel, I., & Kuzle, A. (2019). Problemlösen lernen und unterrichten im SymPa-Projekt: Darstellung einer Projektentwicklung im Rahmen eines Campusschulen-Netzwerkes [Learning and teaching problem solving in the SymPa project: Presentation of a project development within the campus school network]. In K. Kleemann, J. Jennek, & M. Vock (Eds.), *Kooperation von Universität und Schule fördern. Schulen stärken, Lehrerbildung verbessern* (pp. 183–207). Opladen, Berlin, Toronto, Germany: Budrich.

Grouws, D. A. (2003). The teacher's role in teaching mathematics through problem solving. In H. L. Schoen (Ed.), *Teaching mathematics through problem solving: Grades 6–12* (pp. 129–141). Reston, VA: National Council of Teachers of Mathematics.

Gullasch, R. (1967). *Analyse und Synthese bei individuellen Unterschieden in den mathematischen Fähigkeiten untersucht an Schülern 7er Klassen* [Analysis and synthesis of individual differences in mathematical abilities examined in Grade 7 students] (Unpublished doctoral dissertation). Berlin, Pädagogisches Zentralinstitut Berlin, Germany.

Jahn, D. (2014). Durch das praktische Gestalten von didaktischen Designs nützliche Erkenntnisse gewinnen: Eine Einführung in die Gestaltungsforschung [Gaining useful insights through practical educational designs: An introduction to design research]. *Wirtschaft und Erziehung, 66*, 3–15.

Käpnick, F. (1998). *Mathematisch begabte Kinder: Modelle, empirischen Studien und Förderungsprojekte für das Grundschulalter* [Mathematically gifted children: Models, empirical studies and support projects for primary school age]. Frankfurt am Main, Germany: Peter Lang Verlag.

Kilpatrick, J. (1985). A retrospective account of the past twenty-five years of research on teaching mathematical problem solving. In E. A. Silver (Ed.), *Teaching and learning mathematical problem solving: Multiple research perspectives* (pp. 1–16). Hillsdale, NJ: Erlbaum.

Krutetskii, V. A. (1976). *The psychology of mathematical abilities in schoolchildren*. Chicago, IL: The University of Chicago Press.

Kultusministerkonferenz [The Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany]. (2005). *Bildungsstandards im Fach Mathematik für den Primarbereich* [Educational standards in mathematics for the primary school]. Bonn, Germany: KMK.

Kuzle, A. (2017a). Contextual principles for the development of problem solving material. In B. Kaur, W. K. Ho., T. L. Toh, & B. H. Choy (Eds.), *Proceedings of the 41st Conference of the International Group for the Psychology of Mathematics* (Vol. 3, pp. 113–120). Singapore: PME.

Kuzle, A. (2017b). From theory through collaboration into practice: designing a problem solving curriculum for grade 6 students. In T. Dooley & G. Guedudet (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education* (pp. 3849–3856). Dublin, Ireland: DCU Institute of Education and ERME.

Kuzle, A., & Bruder, R. (2016). Probleme lösen lernen im Themenfeld Geometrie [Learning to solve problems in the subject area of geometry]. *mathematik lehren, 196*, 2–8.

Kuzle, A., & Gebel, I. (2016). Development of materials for problem solving instruction in the context of lessons for promoting and improving specific mathematical competences using design based research. In T. Fritzlar, D. Assmuss, K. Bräuning, A. Kuzle, & B. Rott (Eds.), *Problem solving in mathematics education. Proceedings of the 2015 Joint Conference of ProMath and the GDM Working Group on Problem Solving*. *Ars Inveniendi et Dejudicandi 6* (pp. 159–172). Münster, Germany: WTM-Verlag.

Kuzle, A., Gebel, I., & Conradi, C. (2017-2019). *Arbeitsheft für Schülerinnen und Schüler der Klasse 5 und 6 zum Problemlösen* [Student workbook in problem solving for Grade 5 and 6 students]. Potsdam, Germany: Universität Potsdam.

Lester, F. K. (1985). Methodological considerations in research on mathematical problem-solving instruction. In E. A. Silver (Ed.), *Teaching and learning mathematical problem solving: Multiple research perspectives* (pp. 41–69). Hillsdale, NJ: Erlbaum.

Liljedahl, P., Santos-Trigo, M., Malaspina, U., & Bruder, R. (2016). *Problem solving in mathematics education. ICME-13 Topical Surveys*. Cham, Germany: Springer Open.

Lompscher, J. (1975). *Theoretische und experimentelle Untersuchungen zur Entwicklung geistiger Fähigkeiten* [Theoretical and experimental investigations for the development of intellectual ability] (2nd ed.). Berlin, Germany: Volk und Wissen.

National Board of Education. (2004). *National core curriculum for basic education 2004. Finnish National Board of Education*. Vammala, Finland: Vammalan Kirjapaino Oy.

National Board of Education. (2014). *National core curriculum for basic education 2014. Finnish National Board of Education*. Vammala, Finland: Vammalan Kirjapaino Oy.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

Pólya, G. (1973). *How to solve it: A new aspect of mathematical method*. Princeton, NJ: Princeton University Press. (Original work published 1945).

Schoenfeld, A. H. (1985). *Mathematical problem solving*. London, England: Academic Press.

Schunk, D. H. (2008). *Learning theories: An educational perspective* (5th ed.). New York, NY: Prentice Hill.

Senatsverwaltung für Bildung, Jugend und Wissenschaft Berlin, Ministerium für Bildung, Jugend und Sport des Landes Brandenburg. (Eds.). (2015). *Rahmenlehrplan Jahrgangsstufen 1-10. Teil C, Mathematik* [Framework curriculum Grades 1-10. Part C, Mathematics]. Berlin, Potsdam, Germany: Author.

Stanic, G., & Kilpatrick, J. (1989). Historical perspectives on problem solving in mathematics curriculum. In R. I. Charles & E. A. Silver (Eds.), *The teaching and assessing of* *mathematical problem solving* (pp. 1–31). Reston, VA: National Council of Teachers of Mathematics.

The Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 32*(1), 5–8.

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development, 53*(4), 5–23.

Wilson, J. W., Fernandez, M. L., & Hadaway, N. (1993). Mathematical problem solving. In P. S. Wilson (Ed.), *Research ideas for the classroom: High school mathematics* (pp. 57–77). New York, NY: Macmillan.

Wygotski, L. S. (1964). *Denken und Sprechen* [Thinking and speaking]. Berlin, Germany: Akademie-Verlag.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70.

Characteristics of teacher knowledge produced by pre-service mathematics teachers: the case of open-ended problem-based learning

Jani Hannula

University of Helsinki, Finland

One major issue in mathematics teacher education regards the role of university-level mathematics in teacher knowledge. In the context of a design-based research project, an advanced mathematics teacher education course aimed at strengthening the connections between university-level mathematics and school mathematics was developed. In this paper, I present a case study, conducted within the education course, in which I analyse the characteristics of teacher knowledge produced by five small groups of pre-service teachers in an open-ended problem-based learning task. The results indicate the problem-based learning approach has the potential for enhancing specialised content knowledge such as knowledge of different representations of and applications of mathematical concepts. The results also highlight the challenges in using this approach for enhancing horizon content knowledge such as knowledge about the relationships between mathematical concepts. The findings in this case study suggest that problem-based learning can be used to develop mathematics teacher education, although further research is needed to design instructional practices that enhance pre-service teachers' horizon content knowledge.

Keywords
problem-based,
learning,
design-based research,
mathematical knowledge for teaching,
specialised content knowledge,
horizon content knowledge,
pre-service teachers,
teacher knowledge

Correspondence
jani.hannula@helsinki.fi

DOI
https://doi.org/10.31129/LUMAT.7.3.391

1 Introduction

Finnish mathematics teacher education has traditionally put a strong emphasis on subject matter knowledge that is based on courses in advanced mathematics. This tradition is common also in several other countries such as Israel and France (Tatto, Lerman, & Novotna, 2009). The assumption underlying this tradition is that strong subject-matter knowledge rooted in academic mathematics improves teachers' classroom instruction (Even, 2011). However, mathematician Felix Klein (1908/1932) pointed out already more than a century ago that maths teacher education suffers from a 'double discontinuity'. Firstly, when entering university, prospective teachers confront mathematics that is different from what they studied at school. Secondly, after finishing their degrees, novice teachers end up teaching school maths 'traditionally' without any clear connection to the advanced maths they studied at university (Klein, 1908/1932, p. 1).

Nowadays, a large body of maths education research describing different aspects of the double discontinuity exists: the research literature has addressed issues related to secondary–tertiary transition (e.g. Clark & Lovric, 2009), as well as issues related to the development of teacher knowledge (e.g. Moreira & David, 2008). However, the question of utilising knowledge of advanced maths in enhancing mathematical knowledge for teaching has only recently started to gain more attention in the literature, and this research area is still scattered (e.g. Even, 2011; Mosvold & Fauskanger, 2014; Paolucci, 2015; Wasserman, 2016; Zazkis & Leikin, 2010). In particular, there is a lack of research exploring the development of teacher knowledge in course settings aimed at strengthening the connections between advanced maths and school maths (e.g. Wasserman, 2016).

In this paper, I present a case study conducted within a larger design-based research project. The larger project was initiated in order to address the 'second half' of Klein's discontinuity: the aim was to design a maths teacher education course that helps pre-service teachers to connect advanced maths to school maths. The case study presented in this paper focusses on the characteristics of teacher knowledge produced by pre-service teachers in problem-based learning (PBL) task assigned in this course. The purpose of the case study is to add insight into the potentials and challenges presented by the utilisation of open-ended PBL approaches in maths teacher education.

2 Theoretical framework

Theories used and developed in design-based research can be divided into grand theories, orienting frameworks, frameworks for action and domain-specific instructional theories (DiSessa & Cobb, 2004). In the current case study, I use theories from higher education research literature as frameworks for action. Thus, they are utilised in the instructional design of the design artefact. Additionally, domain-specific instructional theories of teacher knowledge are used for the data analysis and problem analysis.

2.1 Frameworks for action

Finnish universities have a long tradition of research-based teacher education (Toom et al., 2010). This research-based approach applies to course contents as well as to teaching methods. On the one hand, course content should be informed by the research literature on the discipline at hand. On the other hand, the teaching and learning methods should be informed by the literature on higher education research

(Toom et al., 2010). During recent decades, social constructivism has been a dominant starting point for instructional design.

In higher education, one widely adopted framework based on social constructivism is the model of constructive alignment (Biggs & Tang, 2011). The idea of constructive alignment is that the intended learning outcomes, and the teaching and learning methods, as well as the assessment, should be carefully designed and aligned with each other. Additionally, teaching and learning methods should be aligned with the theory of social constructivism. Therefore, university students should work actively in a social environment in order to build new knowledge. This kind of approach is stated to raise student engagement and achievement (Figure 1).

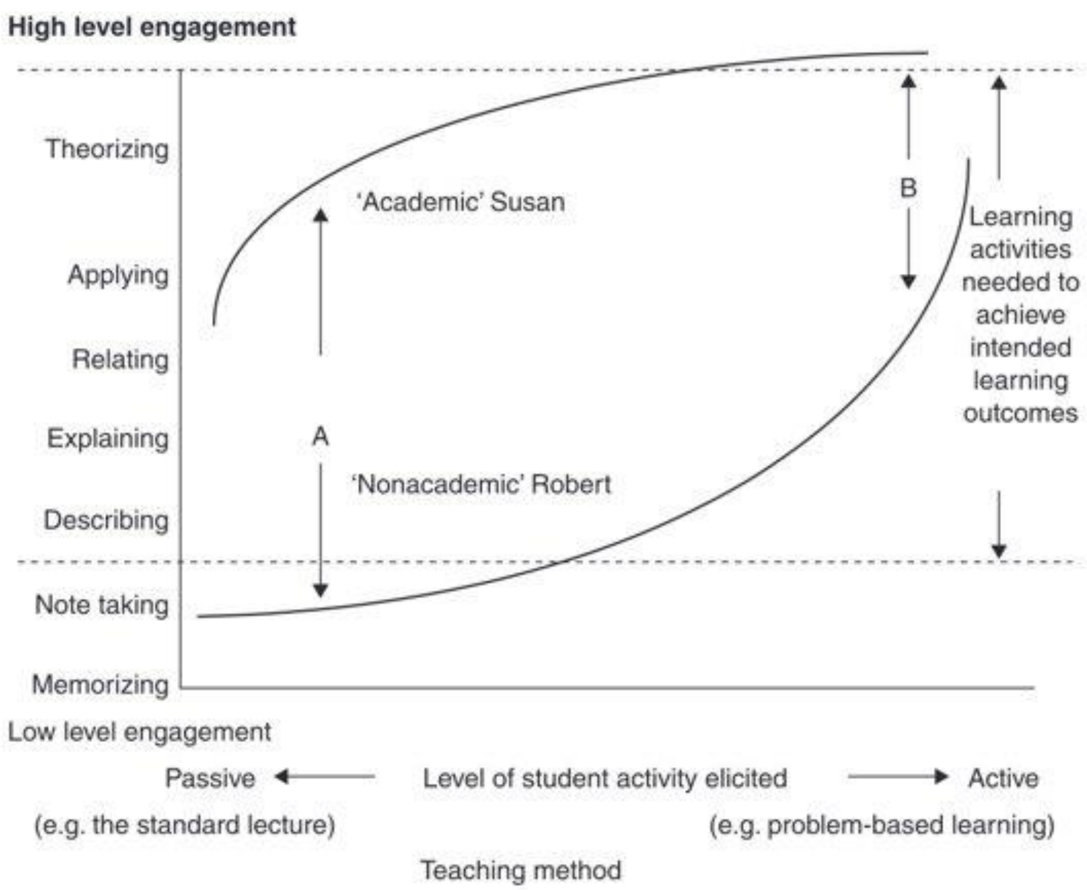


Figure 1. The relationship between teaching method and student engagement (Biggs & Tang, 2011).

With the constructive alignment model, several different teaching and learning methods can be used. In the case study reported in this paper, PBL is adopted. The core idea of PBL is that learning is bound to real-world problems and social interaction. Therefore, PBL aims to enhance not only students' subject matter knowledge but also generic skills such as problem-solving and collaboration (Hmelo-Silver, 2004).

Although there are several ways to implement PBL, it always starts with a real-world problem (a case) analysed by a group of students, and it includes both collaborative and individual work. The seven-step model proposed by Schmidt (1983) (Table 1) is a widely used description of PBL and is also the model adopted in the case study.

Table 1. The steps involved in problem-based learning (Schmidt, 1983).

Step	Description
Step 1	Clarify terms and concepts not readily comprehensible.
Step 2	Define the problem.
Step 3	Analyse the problem.
Step 4	Draw a systematic inventory of the explanations inferred from step 3.
Step 5	Formulate learning objectives.
Step 6	Collect additional information outside the group.
Step 7	Synthesize and test the newly acquired information.

The PBL process begins with a classroom session in which a case is presented to a group(s) of students. The first five steps of Schmidt's model take place in this classroom session. During these steps, students carefully analyse the problem and their previous knowledge and, finally, formulate the learning objectives for the process. That is the group analyses what kind of new knowledge they should learn to solve the problem. It is worth noticing that the cases are typically open-ended and 'ill-structured' (Hmelo-Silver, 2004). Therefore, learning objectives formulated by the groups are typically diverse. After formulating the learning objectives, all the individual students in the group search for literature and theories related to these objectives. The final step of the process is that the group synthesises the new knowledge found during the process.

2.2 Domain-specific instructional theories

Research on teacher knowledge has been greatly influenced by the seminal work of Shulman (1987), whose distinction between subject matter knowledge, pedagogical knowledge and pedagogical content knowledge is still an underlying idea in most of the existing research on the topic. For Shulman, subject matter knowledge refers to 'pure' knowledge of the discipline or the subject (such as mathematics). Pedagogical knowledge refers to general knowledge about the many aspects of pedagogy, such as learning theories. Pedagogical content knowledge, in turn, refers to a 'special amalgam' of content and pedagogy, meaning the special issues related to the teaching and learning of a specific subject.

Shulman's distinction has been elaborated, especially in the mathematical knowledge for teaching (MKT) model (Ball, Thames & Phelps, 2008). In this model, subject matter knowledge (SMK) is divided into common content knowledge (CCK), specialised content knowledge (SCK) and horizon content knowledge (HCK). Similarly, pedagogical content knowledge is divided into the knowledge of content and students (KCS), knowledge of content and teaching (KCT) and knowledge of content and curriculum (KCC). This model aims to give a detailed description of the mathematical knowledge needed in teaching professions (Figure 2).

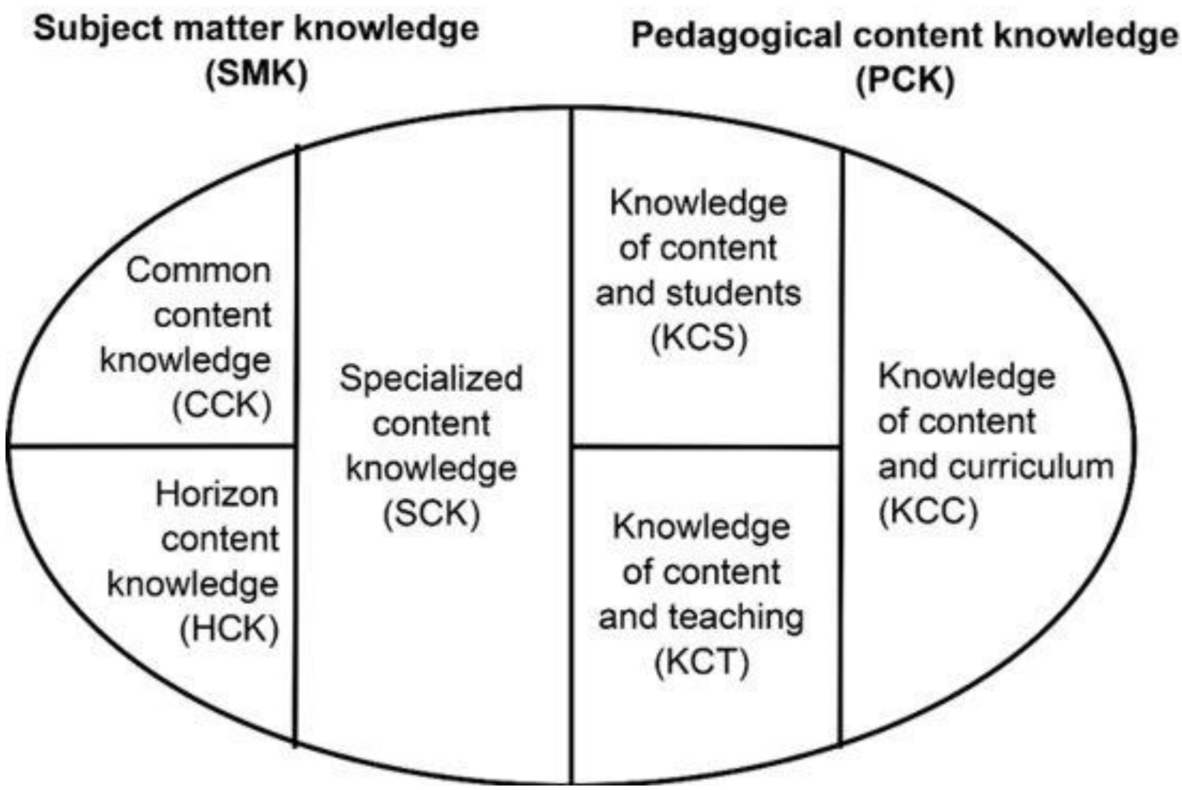


Figure 2. The mathematical knowledge for teaching model (Ball et al., 2008).

Regarding SMK, Ball et al. (2008) separate SCK from CCK by contrasting knowledge needed in teaching and knowledge needed in other professions. CCK is defined as mathematical knowledge that is needed in various professions such as engineering. This kind of knowledge includes, for instance, general mathematical proficiency such as solving equations. SCK, on the other hand, includes knowledge specific for teachers, such as modifying mathematical tasks and being familiar with different representations of the content. These definitions, however, are somewhat problematic as the line between 'common' and 'specialised' seems to be contextual (Carrillo, Climent, Contreras & Muñoz-Catalán, 2013). Lastly, HCK is defined as 'awareness how mathematical topics are related throughout mathematics included in the curriculum' (Ball et al., 2008, p. 403).

With respect to PCK, Ball et al. (2008) state that KCS contains content-specific knowledge of students. This includes, for instance, knowledge of students' typical (mis)conceptions. As another domain Ball et al. delineate is KCT, which is understood as content-specific knowledge regarding classroom orchestration. This kind of knowledge includes, for instance, the ability to sequence lessons 'logically'. Lastly, the domain of KCC is defined as knowledge of curriculum and teaching material such as textbooks (Ball et al., 2008; Koponen, Asikainen, Viholainen & Hirvonen, 2016).

The domains of SCK and HCK are significant in developing maths teacher education. Firstly, these areas are reported as underrepresented in maths teacher education by Finnish in-service teachers and by teacher educators (Koponen et al., 2016). Secondly, the domains are closely related to the second half of Klein's double discontinuity, as they relate to 'specialising' the common content knowledge for teaching purposes and connecting broader disciplinary territory to the school subject and the teaching of it (Jakobsen, Thames & Ribeiro, 2013).

3 Literature review

The core concepts underlying the case study are specialised content knowledge and horizon content knowledge. In the following subsections, I outline prior research related to these concepts from the perspective of curricular and task design as well as practitioners' views and knowledge.

3.1 Specialised content knowledge: Curricular and task design

Recently, new approaches have been proposed to strengthen pre-service teachers' mathematical knowledge for teaching. These approaches typically include tasks that take into account both SMK and PCK (Hoover, Mosvold, Ball & Lai, 2016) and aim to break the boundaries between mathematical content taught in mathematics departments and mathematics pedagogy taught in education departments (e.g. Goos & Bennison, 2018). So far, research on such development has mainly focussed on strengthening elementary and middle school teachers' specialised content knowledge.

Many researchers have proposed task designs that combine the development of SMK and analysis of teaching and/or learning. For instance, Jakobsen, Ribeiro and Mellone (2014) used professional learning tasks in order to reveal prospective primary teachers' mathematical knowledge for teaching. In these tasks, the participants solve mathematical problems and analyse students' answers to the same problem. The results of this study showed that prospective primary teachers' insufficient common content knowledge is problematic in terms of developing SCK. Silver, Clark, Ghousseini, Charalambous and Sealy (2007) used similar professional learning tasks for middle school teachers. They suggest that such an approach helps to 'build or strengthen connections among related mathematical ideas—and to consider these ideas in relation to how students think about the ideas and to a range of pedagogical actions and decisions that affect students' opportunities to learn' (p. 261). Koellner et al. (2007), in turn, present a teaching model called 'problem-solving cycle'. This model was designed for middle school teachers, and it includes solving mathematical problems, lesson planning and analysing the videotaped lessons. As one of their key findings, they argue that the development of specialised content knowledge is evident in the ways the participants compared, reasoned about and made connections between the various solution strategies.

Very few studies concern developmental projects aimed at enhancing secondary teachers' mathematical knowledge for teaching. Typically, mathematics courses given to pre-service secondary maths teachers are based on advanced mathematics and, consequently, may remain unconnected to mathematics taught at school (Moreira & David, 2008). Therefore, some authors (e.g. Papick, 2011; Wasserman, 2016) have proposed tasks for secondary teachers that aim to make advanced mathematical content relevant for developing teacher knowledge. Such tasks aim to combine SMK and authentic classroom situations and to expose the connections between abstract concepts (such as associativity) and school mathematics content (such as mental arithmetic). The results of a study by Wright, Murray and Basu (2016) suggest that such designs can enhance teachers' knowledge of concepts such as inverse elements. However, so far, very little is known of the effects of such course designs and more research is needed to develop such instructional practices (e.g. Wasserman, 2016; Wright et al., 2016).

3.2 Horizon content knowledge: teachers' views and knowledge

During the last decade, the construct of horizon content knowledge has been elaborated upon in response to criticism that it is conceptually problematic (Ball et al., 2008; Jakobsen et al., 2013). HCK is typically associated with knowledge of advanced mathematics, but advanced maths is considered necessary yet not sufficient on its own for the development of HCK (Zazkis & Leikin, 2010). Jakobsen et al. (2013, p. 3128) redefine HCK as 'an orientation to and familiarity with the discipline (or disciplines) that contribute to the teaching of the school subject at hand, providing teachers with a sense for how the content being taught is situated in and connected to the broader disciplinary territory'. In this sense, HCK also includes 'explicit knowledge on ways of and tools for knowledge in the discipline that enables teachers to understand and make judgements of students' statements and reasoning' (p. 3128).

A large proportion of research related to HCK has concentrated on pre-service and in-service teachers' views. In general, both pre-service and in-service teachers have perceived university-level mathematics and school maths as somewhat distinct areas (Hannula, 2018a; Koponen et al., 2016; Mosvold & Fauskanger, 2014; Paolucci, 2015). Additionally, some studies (Mosvold & Fauskanger, 2014; Wasserman, Weber, Villanueva & Mejia-Ramos, 2018) suggest that in-service teachers emphasise the mathematical content at the level they are teaching, disregarding the broader mathematical context. Some pre-service teachers, however, state that advanced mathematics is important for teacher knowledge (Hannula, 2018a; Zazkis & Leikin, 2010). These pre-service teachers, however, do not typically give concrete examples of how advanced mathematical knowledge helps them in their future work. Nevertheless, there is some evidence that pre-service and in-service teachers perceive advanced maths as useful in terms of some specific content such as the history of mathematics and game theory (Even, 2011; Paolucci, 2015).

Some studies examining teachers' knowledge have focussed on the connections between university mathematics and school maths. These studies highlight that exposure to university mathematics does not necessarily change conceptions based on school mathematics and that connecting these domains coherently is difficult for pre-service teachers. For instance, pre-service teachers very often perceive an equation as a process of solving a variable and more rarely perceive it as a statement about the equality of two numbers (Tossavainen, Attorps & Väisänen, 2011). More generally, combining informal reasoning based on graphs and physical interpretations with formal reasoning based on definitions is difficult for pre-service teachers (Viholainen, 2008). Many pre-service teachers explain concepts and properties such as vectors and their distributive property informally using graphs and examples (Hannula, 2018a). These conceptions can become problematic in more complex situations as, for instance, all properties of the sine function cannot be explained using 'triangle trigonometry' (Chin, 2013). As the formal definitions often remain unconnected from informal conceptions, many pre-service teachers hold several misconceptions regarding concepts such as irrational numbers (Sirotic & Zazkis, 2007).

4 Context

4.1 The larger research project

A design-based research project was initiated in 2014 in order to design an advanced course for pre-service mathematics teachers aiming to address the second half of Klein's discontinuity. Following the principles presented by Edelson (2002), an iterative process was composed of an initial problem analysis and case studies of three-course designs (Figure 3).

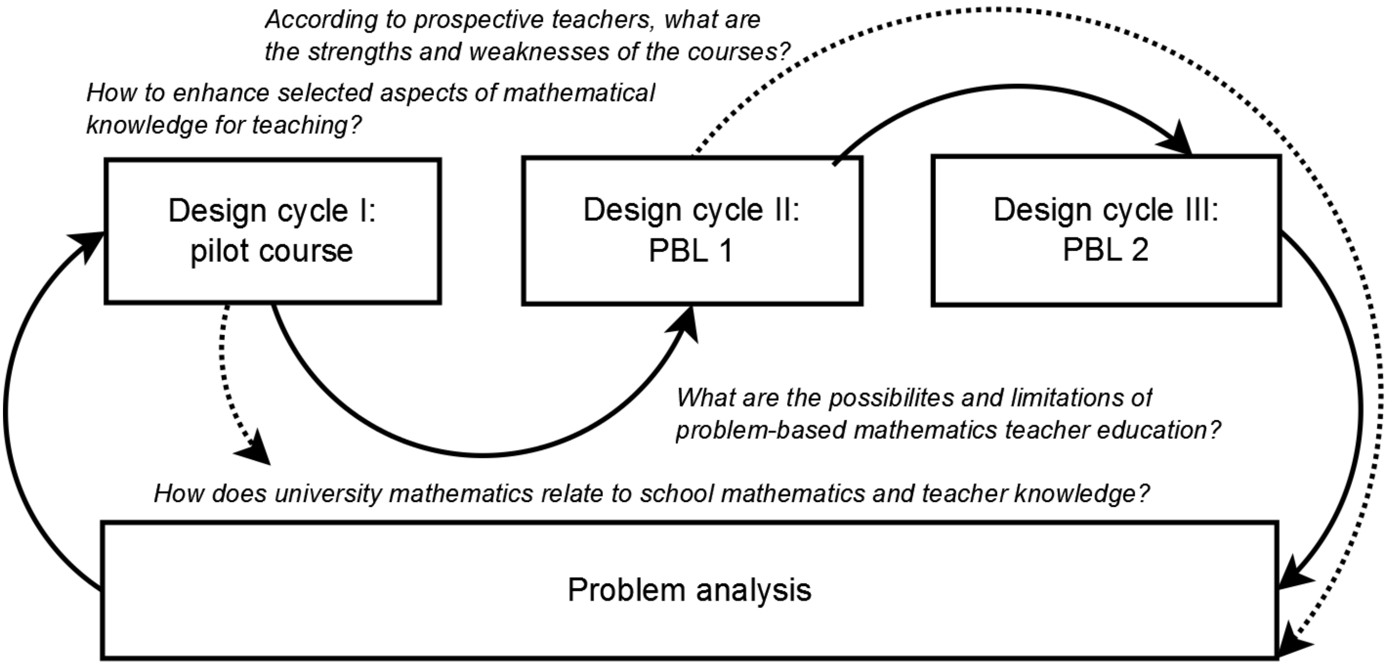


Figure 3. The design process.

Results of the initial empirical problem analysis have been published in Hannula (2018a). Additionally, some results from design cycles I and II have been published in Hannula (2017) and Hannula (2018b), respectively.

4.2 The course design of cycle III

The case study was conducted in the course designed for cycle III. The learning objectives of this course emphasised horizon content knowledge and specialised content knowledge. This seven-week course (3 ECTS credits) had contents related to real analysis, vectors, number systems and logic. The course was taught by the author of this paper.

The course included two PBL tasks as well as lectures and case-based learning. This case study focusses on one of the PBL tasks. This task was implemented following the seven-step model described in section 2.1.

5 Aims and research question

During the first two cycles of the research project, it was notable that the pre-service teachers mainly discussed PCK and SCK in their learning tasks. More accurately, they concentrated, for instance, on forming knowledge related to students' misconceptions or different representations of mathematical content, laying less stress on HCK and certain areas of SCK such as modifying tasks. After the first two cycles, the tasks of the course were refined. During cycle III, the aim was to analyse more closely the characteristics of teacher knowledge produced in one such task. Therefore, in this paper, the following research question is examined:

- What kind of mathematical knowledge for teaching does open-ended PBL provoke in a setting where especially specialised content knowledge and horizon content knowledge are intended to be enhanced?

In relation to prior research literature, the case study has two aims. First, to add to prior literature on SCK and task design by analysing open-ended PBL in the context of pre-service secondary teachers. Second, to give insight into how this task design might enhance the development of HCK, the category that is – in the light of existing literature – problematic from several perspectives.

6 Methods

6.1 Data gathering and study design

The current study focusses on one of the course's PBL tasks. This task was related to vectors and linear algebra. The case presented to pre-service teachers was a hypothetical scenario, formulated by the course teacher, in which a novice teacher ponders the connections between a secondary school course on vectors and a university-level linear algebra course. The original case is written in Finnish as well as an English translation is given in Appendix I.

The pre-service teachers formed six small groups (5–6 participants) in which they worked on the PBL tasks. All of the groups prepared a poster presentation of their work. These poster presentations were then presented to other groups. A more detailed description of the groups' working schedule is given in Appendix II.

The poster presentations of the groups were used as the data in this study. Only the work of five of these groups was analysed. The one presentation left outside the analysis was primarily a comparison of textbooks. That is, their presentation text consisted almost entirely of direct quotations from textbooks and was therefore not considered appropriate for the analysis.

The current study is a case study conducted within a larger project. The course itself is seen as the case. Within the case, five small groups are examined as lower-level units of analysis. Thus, using the terminology of Yin (1994), an embedded case study was conducted (Figure 4).

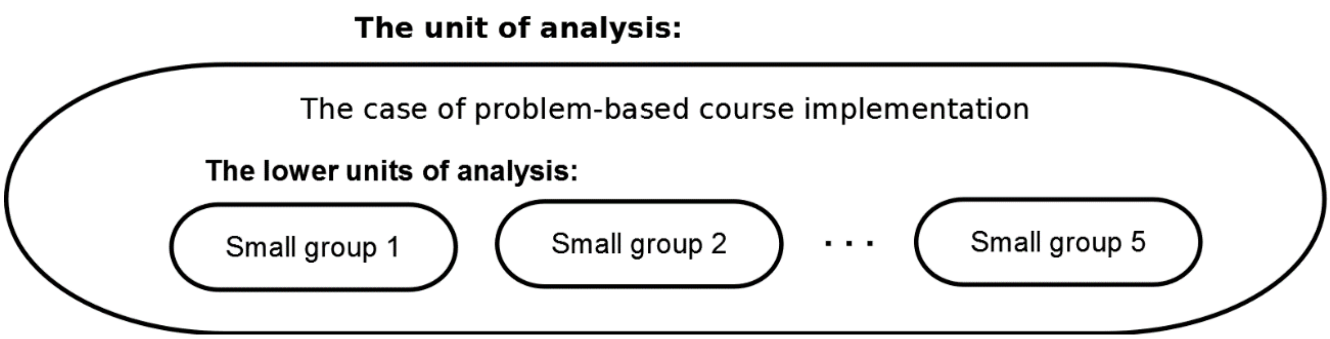


Figure 4. The case study design.

6.2 Participants

The course had 33 participants whom all gave permission to use their work for the research. Of these participants, 29 had mathematics as their major subject, and two had mathematics as a minor subject and were majoring in education. Additionally, two students already had a Master's degree in another subject and were studying intermediate mathematics to qualify as mathematics teachers as well.

6.3 Data analysis

Content analysis (Elo & Kyngäs, 2008) was used as the basic method to examine the data. More accurately, a combination of deductive (directed) and inductive (conventional) content analysis was used. This process was based on the deductive-inductive path model presented by Elo & Kyngäs (2008, p. 110). Two researchers conducted the analysis to enhance the trustworthiness of the process.

First, the units of analysis were determined. A unit of analysis was defined as a written text, a picture or a combination of the two that constitutes one separate idea or statement. The units were initially formed by one researcher, and then the two researchers discussed the outcome and refined the units together. In the second step, the two researchers independently placed each unit of analysis into the categories used in the MKT model.

In the third step, the independent categorisations were cross-tabulated, and Cohen's kappa (Cohen, 1960) was calculated to evaluate the researchers' level of agreement. The kappa value was 0.45, which according to the scale proposed by Landis and Koch (1977) shows that the agreement level was moderate (Table 2).

Table 2. The Kappa Values and Level of Agreement (Landis & Koch, 1977).

Kappa value	Level of agreement
< 0.00	Poor
0.00–0.20	Slight
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Substantial
0.81–1.00	Almost perfect

Looking at the cross-tabulation of the two classifications, it is evident that 87 % of the disagreement is about drawing the line between the following boundaries: CCK vs. SCK, HCK vs. SCK and KCS vs. SCK (Table 3). These three boundary problems are exactly the same that have been pointed out in the literature (e.g. Carrillo et al., 2013).

Table 3. The cross-tabulation of the two researchers' classifications.

		Researcher 2						
		CCK	**SCK**	**HCK**	**KCS**	**KCT**	**KCC**	**Σ**
Researcher 1	**CCK**	12	14	1	0	0	0	27
	SCK	4	41	0	10	0	0	55
	HCK	2	6	2	0	0	0	10
	KCS	0	0	0	12	0	0	12
	KCT	0	0	0	1	0	0	1
	KCC	0	2	0	0	0	4	6
	Σ	18	63	3	23	0	4	111

As the fourth step, contested unit categorisations were discussed by the researchers case by case, and the researchers' justifications for the categorisations were compared. The majority of these units were finally coded as SCK, as the reviewed research literature typically supported this interpretation (Table 4). In the end, a full agreement was achieved. As the last step, the second researcher formed the subcategories. Whenever possible, the subcategories were formed based on the categories suggested in the literature (e.g. Ball et al., 2008; Koponen et al., 2016). For those units of analysis that could not be placed in any of these subcategories, subcategories were formed inductively.

Table 4. An example of the data analysis process.

Example	**Proposed categories**	**Decided category (and justification)**	**Sub-category**
"Using the decomposition representation makes it clear with respect to which basis is the vector given."	SCK and CCK	SCK (recognising what is involved in using a particular representation (Ball et al., 2008))	Characteristics of a representation

7 Results

I present the results of the analysis in subsections for each lower unit of analysis (i.e. by presentation group).

7.1 Group 1

Group 1 focussed on the definitions and properties of mathematical objects as well as on secondary school curricula and textbooks (Table 5).

Table 5. The categorisation of the teacher knowledge produced by Group 1

Common content knowledge (8)	**Specialised content knowledge (2)**	**Horizon content knowledge (1)**	**Knowledge of content and curriculum (4)**
A formal definition (4), properties of a mathematical concept (2), a descriptive definition (1), an example of a concept (1)	Characteristics of representation (1), applications (1)	Relationship of the concepts (1)	Upper secondary school curriculum (2), the content of a textbook (2)

The common content knowledge discussed in this group's work was focussed on the core definitions and properties related to vectors. These definitions and properties were discussed mostly in terms of the definitions given in university-level courses:

> "Set V is a vector space, if it satisfies the following conditions: (...)"

In some parts, the group also connected common content knowledge to knowledge of content and curriculum:

> "For instance, R2 and R3 are vector spaces and in secondary school one typically concentrates on them."

Regarding knowledge of content and curriculum, Group 1 focussed on the upper secondary school curriculum. The group also discussed the content of the curriculum in relation to textbooks:

> "(...) the calculation rules of vectors, unit vector, null vector and inverse vector are presented (...)"

In summary, Group 1 seemed to make an overview of the core mathematical concepts related to the theme ‘vectors’ and connected this knowledge to school maths curricula. However, only occasional observations from the perspective of SCK and HCK were made.

7.2 Group 2

Group 2 focussed on subject matter knowledge from various perspectives (Table 6).

Table 6. The categorisation of the teacher knowledge produced by Group 2.

Common content knowledge (3)	**Specialised content knowledge (6)**	**Horizon content knowledge (2)**
A descriptive definition (1), properties of a mathematical concept (1), alternative definitions (1)	Applications (2), an example (3), linking a representation to an underlying idea (1)	Relationship of the concepts (2)

The group discussed the mathematical content related to dot product both from the perspective of common content knowledge and horizon content knowledge. As an example of common content knowledge, the group presented alternative definitions of the dot product (Figure 5).

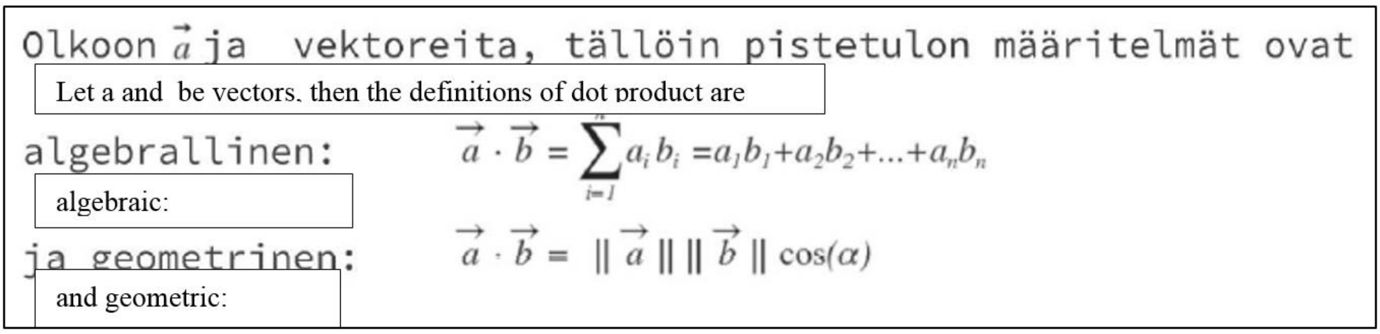


Figure 5. The alternative definitions of dot product presented by Group 2.

Regarding horizon content knowledge, the group explicated the relationship between the concepts:

> “Dot product or scalar product is a real space’s special case of inner product.”

With respect to specialised content knowledge, the group presented applications, examples and a visual representation of dot product:

> "(...) however, if force and transition are not parallel, dot product is used to calculate the work (...)"

In summary, the group discussed subject matter knowledge comprehensively, taking into account the viewpoints related to common content knowledge, to specialised content knowledge and to horizon content knowledge.

7.3 Group 3

Group 3 focussed heavily on knowledge of content and students. Additionally, the group presented some observations related to the different representations (Table 7).

Table 7. The categorisation of the teacher knowledge produced by Group 3.

Specialised content knowledge (2)	**Knowledge of content and students (11)**	**Knowledge of content and teaching (1)**
Linking a representation to an underlying idea (2)	Students' misconceptions (7), difficult content for students (2), students' conceptions (1), students' capability (1)	Choosing a representation in teaching (1)

The knowledge of content and students presented by this group was focussed on students' (mis)conceptions as well as on students' difficulties and capability in terms of mathematical tasks:

> "Students think that dot product gives a vector".

> "The students performed better in adding vectors algebraically than in adding them graphically".

In some parts, the group also discussed the representations of vectors in relation to the underlying ideas and choosing the representation in a teaching situation (Figure 6).

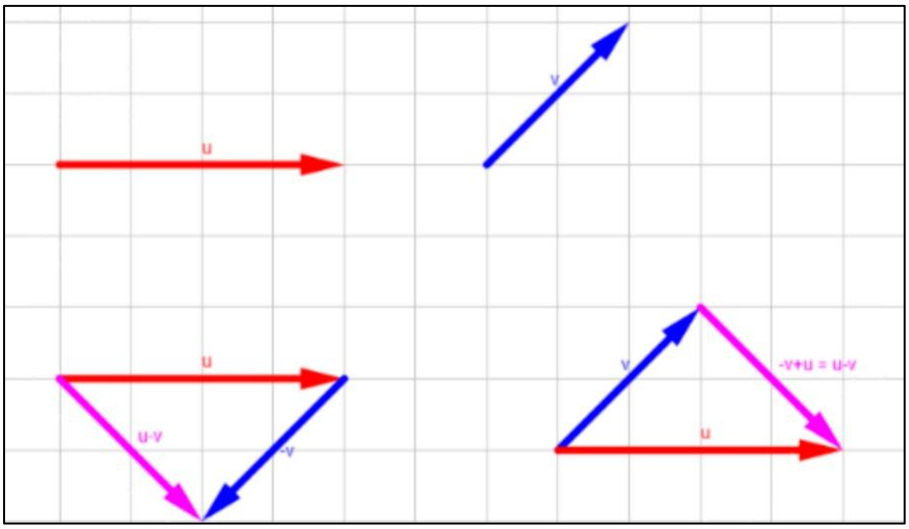


Figure 6. A representation of an inverse vector, the addition of vectors and the commutativity property presented by Group 3.

In summary, Group 3 had a viewpoint on students' conceptions of vectors and operations such as addition and dot product. Some observations of representations were also made, but the overall focus was on summarising students' knowledge and beliefs about vectors.

7.4 Group 4

The most dominant category in the work of Group 4 was specialised content knowledge. This knowledge was focussed entirely on analysing different representations (Table 8).

Table 8. The categorisation of the teacher knowledge produced by Group 4

Common content knowledge (6)	**Specialised content knowledge (25)**	**Knowledge of content and students (2)**	**Knowledge of content and curriculum (1)**
A formal definition (3), mathematical terms (1), a descriptive definition (1), an example of a concept (1)	Characteristics of a representation (20), linking a representation to an underlying idea or other representations (5)	Difficult content for students (2)	Upper secondary school curriculum (1)

To support their analysis of different representations of vectors, the group presented a fair amount of common content knowledge:

> [The decomposition of a vector (title)] "A vector is presented using certain unit vectors".

From the perspective of specialised content knowledge, the group repeatedly analysed what is involved using a certain representation:

> "Using the decomposition representation makes it clear with respect to which basis is the vector given".

To some extent, this group also connected the representations to underlying ideas or other representations such as the idea of representing the vectors of $\mathbb{R}^2$ using unit vectors (1,0) and (0,1) (Figure 7).

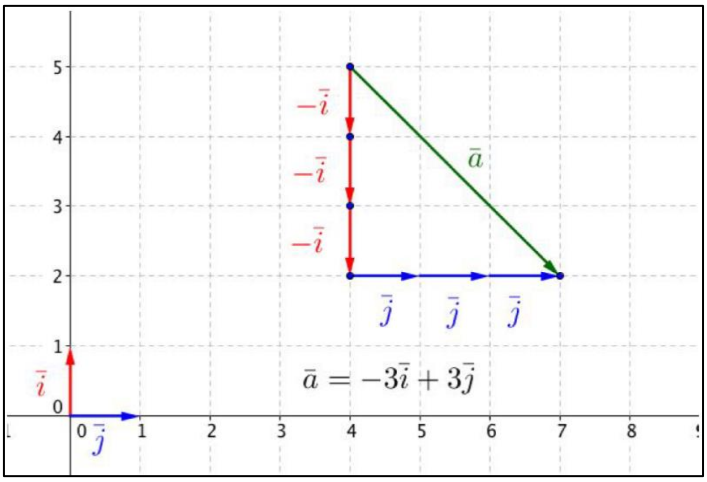


Figure 7. The representation of vector decomposition presented by Group 4.

In summary, Group 4 discussed representations related to vectors with versatility. However, very little attention was given to other areas of teacher knowledge.

7.5 Group 5

Group 5 focussed heavily on the use of vectors in technology and everyday life (Table 9).

Table 9. The categorisation of the teacher knowledge produced by Group 5.

Common content knowledge (3)	Specialised content knowledge (33)	Horizon content knowledge (1)
A formal definition (2), A descriptive definition (1)	Applications (26), Linking a representation to an underlying idea or other representations (3), History of mathematics (2), An example (1), Mathematics and art (1)	Classification of concepts (1)

The group's work was almost entirely composed of the applications of the mathematics related to vectors:

> "Nowadays it [the Bézier curve] is used more extensively in industrial design [...]"
> "With relation to these applications, the group also discussed some other areas of specialised content knowledge such as the history of mathematics:"

With relation to these applications, Group 5 also discussed some other areas of specialised content knowledge such as the history of mathematics:

> "The Bézier curve was developed for designing car bodies by mathematician and engineer Pierre Bézier while working at Renault's car factory in the 1960s."

Common content knowledge received only a little attention in this group's work. However, some definitions related to the applications were presented (Figure 8).

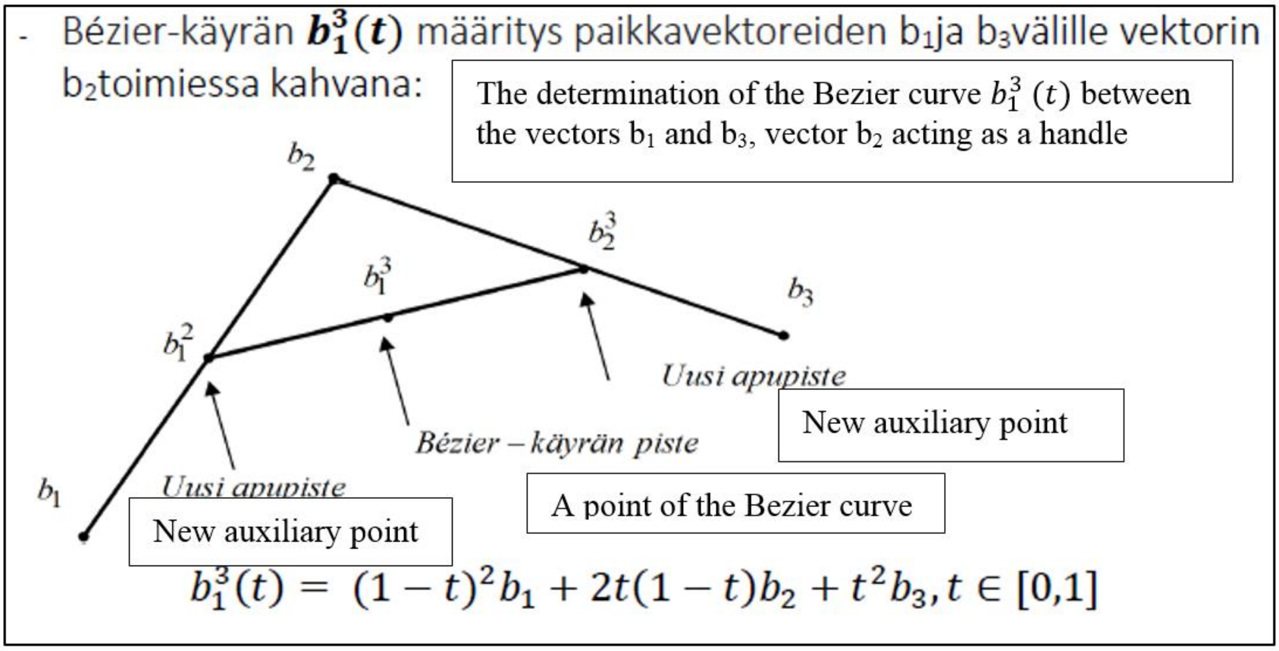


Figure 8. The definition of a Bézier curve presented by Group 5.

In summary, Group 5 presented many applications of vectors. Other areas of subject matter knowledge received only a little attention.

7.6 Summary

Overall, only Group 3 emphasised PCK whereas other groups focussed on SMK. Looking at the other four groups, the Group 1 focussed on common content knowledge and knowledge of content and curricula. However, Groups 2, 4 and 5 produced a considerable amount of specialised content knowledge (Table 10).

Table 10. The categorisation of the teacher knowledge produced by Groups 1, 2, 3, 4 and 5.

	Subject matter knowledge			*Pedagogical content knowledge*		
	CCK	**SCK**	**HCK**	**KCS**	**KCT**	**KCC**
Group 1	8	2	1	0	0	4
Group 2	3	6	2	0	0	0
Group 3	0	2	0	11	1	0
Group 4	6	25	0	2	0	1
Group 5	3	33	1	0	0	0

Groups 2, 4 and 5 produced qualitatively varying kinds of specialised content knowledge. Group 5 presented various applications of mathematics in everyday life and technology. However, these applications were rarely connected to common content knowledge or school mathematics contents. Group 4 focused on analysing characteristics of different representations of mathematical content. These representations were, however, rarely explicitly connected to other domains such as the structure of the mathematical theory. In contrast, Group 2 connected all the subdomains of SMK and, thereby, presented a wide overall view of one mathematical concept.

8 Discussion

In relation to the research question, the current study showed that open-ended PBL provoked the production mainly of specialised content knowledge. Compared to knowledge production seen in previous design cycles (Hannula, 2017; Hannula 2018b) a shift in emphasis from PCK to SMK seems evident. Thus, the proposed PBL approach seems to have the potential for enhancing SCK in mathematics teacher education. However, somewhat in conflict with the learning objectives, the category of HCK received only a little attention overall and only one group included HCK more notably in their work.

In terms of SCK, the current study adds to the prior research by describing secondary teachers' different approaches to the PBL task. The results of the current study show that open-ended problem-solving tasks may provoke the development of different subareas of SCK such as applications and representations. Thus, the design of the current study contrasts such designs proposed in prior research that support the development of narrower areas of SCK such as comparing solution strategies (e.g. Koellner et al., 2007). It may be that the development of teacher knowledge through PBL will consistently be narrower than that achieved with designs in which different components of teacher knowledge are more explicitly involved in the task (e.g. Silver et al., 2007). In terms of the development of instructional design, the challenge is how to avoid encouraging a fragmented view of knowledge and to advance the synthesis of the viewpoints of different student-teacher groups.

The results showed that this PBL task provoked relatively little consideration with respect to HCK. The existing research literature indicates that both pre-service and in-service teachers view advanced mathematical knowledge as being of only limited significance for teacher knowledge (Even, 2011; Hannula, 2018; Mosvold & Fauskanger, 2014; Paolucci, 2015). In this study, it was found that a similar trend seems evident also in authentic learning situations. However, as the category of HCK is conceptually problematic, it might be that the pre-service teachers' views on the development of HCK in the task differ from the observations presented in this paper. Therefore, it would be valuable to conduct further research on pre-service teachers' conceptions of the development of HCK in these tasks. Additionally, the disregard of HCK might be caused by the difficulties in finding coherent connections between university mathematics and school mathematics (Chin, 2013; Hannula, 2018). Further studies could examine pre-service teachers' conceptions of vectors and linear algebra and find ways to support their HCK related to these topics. In the future, the cases could be designed more in line with the notion of HCK. Such cases should encourage utilising the knowledge of advanced mathematics in classroom situations such as using the knowledge of proof by contradiction to understand students' reasoning.

9 Limitations and conclusion

Some limitations of the current case study must be taken into account. Firstly, participants' knowledge or conceptions were not systematically tested by using standardised procedures such as pre- and postintervention by using questionnaires. Instead, the study provided information on an authentic learning situation. Secondly, case studies can only provide contextual information on the learning process. Therefore, the results of the current study cannot be generalised to other contexts. However, by means of a careful report, a repeatable procedure and a transferable design have been provided for further study and development. Lastly, this case study supports the view that the categories of the MKT model are conceptually problematic, as they seem to be contextual and not mutually exclusive. Consequently, further research is needed to clarify the components of mathematical knowledge for teaching.

Overall, as several groups produced diverse outputs categorised in specialised content knowledge, the findings of the current study suggest that the PBL approach enables pre-service secondary teachers to specialise their content knowledge for teaching purposes. Such knowledge is underrepresented in current maths teacher education (Koponen et al., 2016) and is associated with improved teaching practices (Hoover et al., 2016). Therefore, as the open-ended PBL approach provoked the development SCK, it seems to be one promising approach for developing maths teacher education. However, further research from different standpoints such as pre-service teachers' views and knowledge is needed to evaluate the overall relevance and effect of such an approach. Additionally, as the amount of HCK produced in the task was relatively small, the findings of the current study support prior research literature (e.g. Wasserman, 2016; Wright et al., 2016) in the recommendation that new instructional practices and practice-based research are needed for the development of HCK.

Acknowledgements

This research was supported by the Jenny and Antti Wihuri Foundation under grant numbers 00160077 and 00170067. I would like to express my gratitude to Juulia Lahdenperä who kindly helped me in the data analysis. I would also like to thank the members of our research group for commenting on an early version of the manuscript.

References

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389–407.

Biggs & Tang (2011). *Teaching for quality learning at university* (4th ed.). Maidenhead: Open University Press.

Carrillo, J., Climent, N., Contreras, L., & Muñoz-Catalán, M. (2013). Determining specialised knowledge for mathematics teaching. In B. Ubuz, C. Haser, & M. A. Mariotti (Eds.), *Proceedings of the Eighth Congress of the European Society for Research in Mathematics Education* (pp. 2985–2994). Ankara, Turkey: Middle East Technical University and ERME.

Chin, K. E. (2013). *Making sense of mathematics: Supportive and problematic conceptions with special reference to trigonometry*. (PhD thesis) The University of Warwick.

Clark & Lovric (2009). Understanding secondary–tertiary transition in mathematics. *International Journal of Mathematical Education in Science and Technology, 40*(6), 755–776.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46.

DiSessa, A. A., & Cobb, P. (2004). Ontological innovation and the role of theory in design experiments. *The Journal of the Learning Sciences, 13*(1), 77–103.

Edelson, D. C. (2002). Design research: What we learn when we engage in design. *The Journal of the Learning Sciences, 11*(1), 105–121.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing, 62*(1), 107–115.

Even, R. (2011). The relevance of advanced mathematics studies to expertise in secondary school mathematics teaching: Practitioners' views. *Zdm Mathematics education, 43*(6-7), 941–950.

Goos, M., & Bennison, A. (2018). Boundary crossing and brokering between disciplines in pre-service mathematics teacher education. *Mathematics Education Research Journal, 30*(3), 255–275.

Hannula, J. (2017). Subject matter knowledge and pedagogical content knowledge in the learning diaries of prospective mathematics teachers. In T. Dooley & G. Gueudet (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education* (pp. 3312–3319). Dublin, Ireland: DCU Institute of Education & ERME.

Hannula, J. (2018a). The gap between school mathematics and university mathematics: prospective mathematics teachers' conceptions and mathematical thinking. *Nordic studies in mathematics education 23*(1), 67–90.

Hannula, J. (2018b). Matematiikan aineenopettajaopiskelijoiden matemaattinen ja pedagoginen sisältötieto ongelmalähtöisessä oppimisessa. [Pre-service mathematics teachers' mathematical and pedagogical content knowledge in problem-based learning.] *FMSERA Journal, 2*(1), 134–144

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review, 16*(3), 235–266.

Hoover, M., Mosvold, R., Ball, D. L., & Lai, Y. (2016). Making progress on mathematical knowledge for teaching. *The Mathematics Enthusiast, 13*(1/2), 3–34.

Jakobsen, A, Ribeiro, C. M., & Mellone, M. (2014). Norwegian prospective teachers' MKT when interpreting pupils' productions on a fraction task. *Nordic Studies in Mathematics Education, 19*(3-4), 135–150.

Jakobsen, A., Thames, M. H., & Ribeiro, C. M. (2013). Delineating issues related to horizon content knowledge for mathematics teaching. In B. Ubuz, C. Haser, & M.A. Mariotti (Eds.), *Proceedings of the Eighth Congress of the European Society for Research in Mathematics Education* (pp. 1389–1398). Ankara, Turkey: Middle East Technical University and ERME.

Klein, F. (1932). *Elementary mathematics from an advanced standpoint: Arithmetic, algebra, analysis*. New York: Dover Publications. (Original work published in 1908).

Koellner, K., Jacobs, J., Borko, H., Schneider, C., Pittman, M. E., Eiteljorg, E . . . Frykholm, J. (2007). The problem-solving cycle: A model to support the development of teachers' professional knowledge. *Mathematical Thinking and Learning, 9*(3), 273–303.

Koponen, M., Asikainen, M. A., Viholainen, A., & Hirvonen, P. E. (2016). Teachers and their educators' views on contents and their development needs in mathematics teacher education. *The Mathematics Enthusiast, 13*(1/2), 149–170.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.

Moreira, P. C., & David, M. M. (2008). Academic mathematics and mathematical knowledge needed in school teaching practice: Some conflicting elements. *Journal of Mathematics Teacher Education, 11*(1), 23–40.

Mosvold, R., & Fauskanger, J. (2014). Teachers' beliefs about mathematical horizon content knowledge. *International Journal for Mathematics Teaching and Learning, 9*(3), 311–327.

Paolucci, C. (2015). Changing perspectives: Examining the potential for advanced mathematical studies to influence pre-service teachers' beliefs about mathematics. *Teaching and Teacher Education, 49*, 97–107.

Papick, I. J. (2011). Strengthening the mathematical content knowledge of middle and secondary mathematics teachers. *Notices of the AMS, 58*(3), 389–392.

Schmidt, H. G. (1983). Problem-based learning: Rationale and description. *Medical education, 17*(1), 11–26.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1–23.

Silver, E. A., Clark, L. M., Ghousseini, H. N., Charalambous, C. Y., & Sealy, J. T. (2007). Where is the mathematics? Examining teachers' mathematical learning opportunities in practice-based professional learning tasks. *Journal of Mathematics Teacher Education, 10*(4-6), 261–277.

Sirotic, N., & Zazkis, A. (2007). Irrational numbers: The gap between formal and intuitive knowledge. *Educational Studies in Mathematics, 65*(1), 49–76.

Tatto, M. T., Lerman, S., & Novotna, J. (2009). Overview of teacher education systems around the world. In R. Even, & D. L. Ball (Eds.), *The professional education and development of teachers of mathematics: The 15th ICMI study* (pp. 15–24). New York: Springer.

Toom, A., Kynäslahti, H., Krokfors, L., Jyrhämä, R., Byman, R., Stenberg, K. . . . Kansanen, P. (2010). Experiences of a research-based approach to teacher education: Suggestions for future policies. *European Journal of Education, 45*(2), 331–344.

Tossavainen, T., Attorps, I., & Väisänen, P. (2011). On mathematics students' understanding of the equation concept. *Far East Journal of Mathematical Education, 6*(2), 127–147.

Viholainen, A. (2008). Finnish mathematics teacher student's informal and formal arguing skills in the case of derivative. *Nordic Studies in Mathematics Education, 13*(2), 71–92.

Wasserman, N. H. (2016). Abstract algebra for algebra teaching: Influencing school mathematics instruction. *Canadian Journal of Science, Mathematics and Technology Education, 16*(1), 28–47.

Wasserman, N., Weber, K., Villanueva, M., & Mejia-Ramos, J. P. (2018). Mathematics teachers' views about the limited utility of real analysis: A transport model hypothesis. *The Journal of Mathematical Behavior*, *50*, 74–89.

Wright, M., Murray, E., & Basu, D. (2016). Impact of abstract algebra knowledge on the teaching and learning of secondary mathematics. In T. Fukawa-Connelly, N. Engelke Infante, M. Wawro & S. Brown (Eds.), *Proceedings of 19th Annual Conference on Research in Undergraduate Mathematics Education* (pp. 481–495). Pittsburgh, Pennsylvania.

Yin, R. K. (1994). *Case study research: Design and methods* (2nd ed.). Newbury Park, CA: SAGE Publications.

Zazkis, R., & Leikin, R. (2010). Advanced mathematical knowledge in teaching practice: Perceptions of secondary mathematics teachers. *Mathematical Thinking and Learning, 12*(4), 263–281.

Appendix I: The case of the PBL project

The original case written in Finnish:

Antti Aineenopettaja pääsi matematiikan ja tilastotieteen laitokselle opiskelemaan matematiikan aineenopettajaksi, mikä oli hänelle mieluisa vaihtoehto, sillä koki oppineensa matematiikkaa koulussa hyvin ahkeran opiskelunsa myötä (matematiikka oli hänen vahvin kouluaineensa, yleensä 9 tai 10). Lisäksi opetustyö tuntui hänestä hyvältä uravalinnalta; hän piti jo koulussa siitä, kun sai auttaa muita oppilaita tehtävissä.

Yliopistossa yhtenä ensimmäisistä kursseistaan hän opiskeli kurssin ”Lineaarialgebra ja matriisilaskenta”. Kurssilla oli lähtötasotesti, joka käsitteli vektoreita. Näin ollen Antti päätteli, että kurssi tulisi liittymään jollakin tavalla lukiossa opetettuihin vektoreihin ja hän kertasikin nopeasti pääkohdat lukion vektorikurssista.

Antti käsitti vektorit lukion pohjalta nuolina eli olioina, joilla on suunta ja suuruus. Hyvänä esimerkkinä käytännön esimerkkinä hän muisti joen virtauksen, joka vie venettä tiettyyn suuntaan tietyllä voimalla. Tällaisia kutsuttiin vektorisuureiksi, joita oli mukava laskea. Erityisesti kun niitä tarkasteltiin koordinaatistossa, jossa vektori esitettiin esimerkiksi muodossa 3i + 4j. Vektoreiden yhteenlaskun ja vakiolla kertomisen Antti ymmärsi hyvin. Yksi oudompi laskutoimitus oli pistetulo, joka määriteltiin kaavalla $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}|\cdot|\mathbf{b}|\cdot \cos \alpha$, missä α on vektoreiden **a** ja **b** välinen kulma. Pistetulosta Antti oppi, että kahden vektorin pistetulo on nolla silloin, kun vektorit ovat toisiaan vasten kohtisuorassa.

Lineaarialgebran kurssilla käsiteltiin myös vektoreita, mutta nyt vektorit esitettiin järjestettynä parina, esim. (3, 6) tai jonona, esim. (2,-3,4,1,-3). Kahden vektorin pistetulo määriteltiin kaavalla, $\mathbf{a} \cdot \mathbf{b} = a_1 \cdot b_1 + a_2 \cdot b_2 + \ldots + a_n \cdot b_n$, eli laskemalla vektoreiden komponenttien tulot ja laskemalla ne yhteen. Vektorit olivat Antin mielestä tällä kurssilla jotenkin erilaisia kuin lukiossa; lukiossa esimerkiksi pisteen (3,6) paikkavektoriksi sanottiin vektoria 3i + 6j, nyt itse piste (3,6) oli vektori. Pian edettiin matriiseihin, jotka tuntuivat taas Antista harppaukselta johonkin uuteen. Lineaarialgebran kurssi meni lopulta Antilla hyvin ja hän koki oppineensa asioita, joskin kokonaisuus jäi osittain hajanaiseksi.

Myöhemmin opinnoissaan Antti törmäsi myös siihen, miten vektori voidaan määritellä tarkasti myös geometrisesti keskenään yhtenevien suuntajanojen ekvivalenssiluokkana. Nyt Antilla oli useita eri tapoja lähestyä vektoreita ja kokonaisuus alkoi hahmottua.

Opetusharjoittelussa Antti pääsi opettamaan lukion vektorikurssia. Hän kävi läpi oppikirjaa, jossa aluksi vektoreita esitettiin yleisesti geometrisina otuksina ja fysikaalisten sovellusten kautta. Pian siirryttiin laskemaan vektoreilla koordinaatistossa. Antin oli määrä pitää oppitunteja liittyen mm. vektorin käsitteeseen, vektoreiden laskusääntöihin, vektoreiden esittämiseen kantavektoreiden avulla ja pistetuloon.

Antti mietti, että hänellä on kokonaisuus hallussa kohtuullisen hyvin, mutta vieläkin jotkut asiat olivat vähän epäselviä. Hän mietti mm., miten ”geometriset vektorit” ja ”koordinaatistovektorit” liittyvät toisiinsa, mikä pistetulo oikeastaan on, käytetäänkö lineaarialgebraa jossakin muussakin kuin fysiikassa ja mihin matematiikan aloihin lineaarialgebra oikeastaan liittyy. Hän tiesi, että hän selviäisi opetusharjoittelusta, vaikkei osaisikaan vastata tarkasti edellisiin kysymyksiin, mutta olo tuntui silti hieman epävarmalta: kokonaisuuden hahmottamisessa oli vielä aukkoja!

An English translation of the case:

Antti got a right to study as mathematics teacher at a Department of Mathematics and Statistics. This was a pleasant choice for him, as he thought that he had learned maths well in school due to his studious attitude. (Mathematics was his strongest subject, usually the grade was 9 or 10 out of 10.) Additionally, teacher’s work seemed as a good career choice for him. He enjoyed helping other students already in school.

One of his first courses at the university was “Liner algebra and matrices”. The course had a placement test regarding vectors. Therefore, Antti concluded that the course had something to do with vectors learned at secondary school. He revised the main topics of the secondary school vector course.

From his secondary school experiences, Antti perceived vectors as arrows i.e. objects that have a direction and a magnitude. As a good example to him, was the flow of river that takes the boat into a certain direction in a certain force. These were called as vector magnitudes and Antti enjoyed calculating these. Especially, when the vectors were examined in a set of coordinates and they were expressed for instance in form 3i + 4j. Antti understood well the addition and multiplication of vectors. One of the more complicated operations was the dot product that was defined with the formula $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}| \cdot |\mathbf{b}| \cdot \cos \alpha$, where α is the angle between the vectors **a** and **b.** Antti learned that the dot product of two vectors is zero if the vectors are perpendicular to each other.

Vectors were discussed also in the linear algebra course, but now the vectors were expressed as ordered pairs such as (3, 6) or sequences such as (2,-3,4,1,-3). The dot product was defined with a formula $\mathbf{a} \cdot \mathbf{b} = a_1 \cdot b_1 + a_2 \cdot b_2 + \ldots + a_n \cdot b_n$, that is, calculating the multiplications of the components of the vectors and adding them up. In this course, Antti thought that vectors were different from the ones in secondary school. In secondary school, the place vector of the point (3, 6) was $3i + 6j$ and now the point itself was a vector. Soon the course proceeded to matrices and to Antti this was again something different. In the end, the course went well and Antti thought that he had learned a great deal, even though the big picture remained a bit fuzzy.

Later in his studies, Antti run into a geometric definition of vectors through equivalence classes of directed line segments. Now Antti had several approaches to vectors and the big picture started to take shape.

During his practical training, Antti has to teach the vector course of secondary school. He went through the textbook that started by introducing the vectors as geometric object and through physics applications. Soon the textbook proceeded to calculations in a set of coordinates. Antti was supposed to give lessons regarding the vector concept, the operations of vectors, the expression of vectors through basis vectors and dot product.

Antti thought that he handled these topics quite well, even though some things were still unclear to him. He wondered things such as “How are ‘geometric vectors’ related to ‘coordinative vectors’?”, “What is dot product, really?”, “Is linear algebra applied in any other discipline than physics” and “Which areas of mathematics is linear algebra related to?”. He knew that he would survive his practical training without having the answers but he felt a bit uncertain: the big picture is was not yet without gaps!

Appendix II: The working schedule of the groups

Week	The objective	The steps of Schmidt's model	Instruction / individual work
Week 1	Analysing the case and formulating learning objectives	1-5	2 x 45 minutes of instruction
Week 2	Collecting additional information outside the group	6	Individual work
Week 3	Making the synthesis and presentation of the work	7	Individual work + 2 x 45 minutes of instruction

Pedagogical experiments with MathCheck in university teaching

Terhi Kaarakka[1], Kirsi Helkala[2], Antti Valmari[3] and Marjukka Joutsenlahti

[1] Tampere University, Faculty of Information Technology and Communication Sciences/Computing Sciences, Finland
[2] Norwegian Defence University College, Norwegian Defence Cyber Academy, Norway
[3] University of Jyväskylä, Faculty of Information Technology, Finland

MathCheck is a relatively new online tool that gives students feedback on their solutions to elementary university mathematics and theoretical computer science exercises. MathCheck was designed with constructivism learning theory in mind and it differs from other online tools as it checks the solutions step by step and shows a counter-example if the step is incorrect. It has been in student use since the autumn of 2015 and under design-based research from the first online day. The main research questions of this study are the following. 1) How can the usage of MathCheck support the aspects of conceptual understanding and procedural fluency of constructivism learning? 2) How can MathCheck empower both students and teachers in the education of mathematics? This paper presents the results of five pedagogical experiments considering both students' and teachers' point of views. In each experiment, the students have suggested improvements, which have affected the further development of MathCheck. In general, both students and teachers have given positive feedback on MathCheck. MathCheck seems to support learning better than tools that only provide the "incorrect"/"correct" verdict after checking the answer. MathCheck is suitable for independent studying as well as an addition to traditional lectures. In the best case, it can reduce teachers' workload during courses.

Keywords
MathCheck, pedagogy, mathematics, education

Correspondence
terhi.kaarakka@tuni.fi

DOI
https://doi.org/10.31129/LUMAT.7.3.428

1 Introduction

Traditionally university mathematics has been taught with the pencil and paper method. Over the last decade, computers and online tools for mathematics have established their place as a part of mathematics courses (Mäkelä, 2016). There are plenty of online tools for students to use in mathematics. One popular type of online mathematics tools simplify expressions, evaluates expressions, and solves equations. Examples of such tools are Matlab (MathWorks), Wolfram Alpha (Wolfram Alpha) and GeoGebra (GeoGebra). The latter is more used in upper secondary schools while Matlab and Wolfram Alpha are more used in universities. Such tools are convenient when the student already understands the mathematics behind the operation. However, teachers have observed that students are using these tools more often just to get correct answers without understanding mathematics.

A prime example of another popular type of online tools for mathematics education is STACK (System for Teaching and Assessment using a Computer algebra Kernel) (Sangwin, 2015). With it, the teacher provides problems for the student (Mäkelä et al., 2016). The student solves each problem with a pencil and paper (at least the teacher hopes so) and then types the final answer to the website. Let us consider the simplification of $\sqrt{3a + a^2 + 1} - a$ as an example. The student computes $\sqrt{3a + a^2 + 1 - a} = \sqrt{a^2 + 2a + 1} = \sqrt{(a+1)^2} = a + 1$ on paper and types $a + 1$ to STACK. STACK checks the answer immediately and gives feedback telling that the answer is incorrect. STACK compares the student's answer to the teacher's answer with Maxima (a symbolic algebra system) (Maxima) and reports whether or not they are mathematically equivalent. STACK does not tell in its feedback where the possible mistake has happened – it cannot, because it has only been given the final answer and not the intermediate steps that led to it. This software is used in many universities and its pedagogical utility has been the subject of much research and discussion in different perspectives (Mäkelä, 2016), (Pelkola, Rasila, Sangwin, 2018).

Unfortunately, students can use these systems to support behavioural learning. While it is possible for a teacher to build in STACK task sets and feedback systems that also ensure in-depth learning, this requires a lot of teacher work. Therefore, it is possible that both of these methods support behavioural learning, where the aim is on the right answers, and the wrong answers are disregarded. MathCheck differs from the tools mentioned above as it gives feedback on all steps of the solution that the student types, not on just the final answer. As the feedback on an incorrect step, it gives a counter-example. Therefore, MathCheck could support constructivism learning, as in constructivism learning the learner builds her knowledge and concept understanding by making sense of all information perceived from her experiences (Bada, 2015).

In this paper, we study the usage of MathCheck in teaching finding out answers to the following questions. How can the usage of MathCheck support the aspects of conceptual understanding and procedural fluency of constructivism learning? How can MathCheck empower both students and teachers in the education of mathematics?

2 MathCheck as a constructivism learning platform

Nowadays students' role in the learning process is emphasised. The constructivism learning process gives students the responsibility of learning (Weegar & Pacis, 2012). A student must oneself be an active thinker and processor, and construct new information on top of old information. With a suitable learning process, it is possible to affect different areas of mathematics learning. As a theoretical framework to describe mathematics learning, we use the concept of mathematical proficiency, which consists of the following five components (National Research Council, 2001)

1. *conceptual understanding*: comprehension of mathematical concepts, operations, and relations,
2. *procedural fluency*: skill in carrying out procedures flexibly, accurately, efficiently and appropriately,
3. *strategic competence*: ability to formulate, represent, and solve mathematical problems,
4. *adaptive reasoning*: a capacity for logical thought, recreation, explanation, and justification and
5. *productive disposition*: a habitual inclination to see mathematics as sensible, useful, and worthwhile, coupled with a belief in diligence and ones' own efficacy.

Of the components listed above, MathCheck (MathCheck; Valmari & Kaarakka, 2016; Valmari & Rantala, 2019) aims to support especially *conceptual understanding and procedural fluency*. MathCheck supports these from the constructivism learning point of view as it gives feedback on all steps of the solution that the student types, not on just the final answer. As the feedback on an incorrect step, it gives a counter-example. The current version of MathCheck also draws graphs of the expressions on both sides of the error place (this feature had not been implemented yet in the versions that were used in the experiments reported in this study). These give the student a starting point for tracing the error. When a student's erroneous thought chain is overturned by a counterexample, the student must rethink his preconceptions. Then the student rebuilds his reasoning and this is close to radical constructivism. At the same time, we are working in the students' (Vygotsky) zone of proximal development. Boudourides nicely explores various sub-categories of constructivism and explains Vygotsky's theory in his article Constructivism, Education, Science, and Technology (Boudourides, 2003).

Figure 1 shows the feedback that MathCheck gives on our example. It is clear that the step $\sqrt{(a+1)^2} = a + 1$ is incorrect. The green and red graphs show that for negative enough numbers, $a + 1$ yields negative values while $\sqrt{(a+1)^2}$ yields positive values. The correct step is $\sqrt{(a+1)^2} = |a + 1|$.

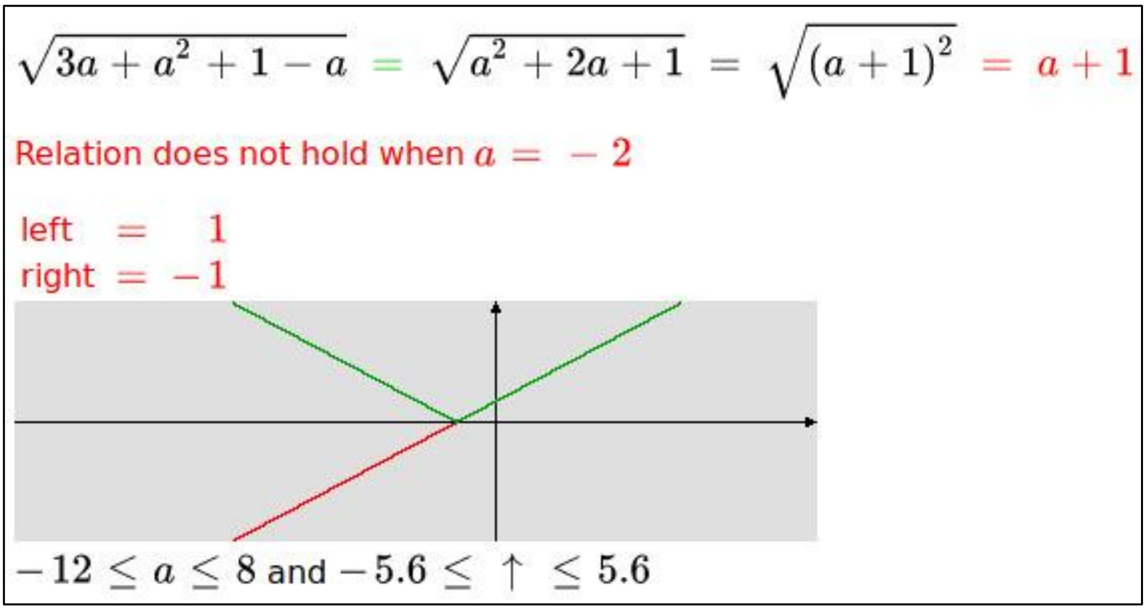


Figure 1. An example of how MathCheck shows errors.

To produce the feedback in Figure 1, MathCheck needs absolutely no model solution or other contribution by the teacher. It suffices that the student types the solution to the main page of MathCheck and presses the submit button. The default mode of MathCheck works by checking the mathematical correctness of each equality and inequality in the input, without assuming that the input should be an answer to some specific problem or that the computation in the input should follow some pre-specified path.

When checking a relation in the simplification mode, MathCheck first tries to prove it correct. If that succeeds, MathCheck shows the relation symbol in green. The proof engine of MathCheck is rather straightforward, but also weak. If MathCheck fails to prove the relation, it tries to find a counter-example by trying many combinations of values of the variables in question. If MathCheck finds a counter-example, it prints the relation symbol, the expression to its right, and the counter-example in red. Otherwise, it prints the relation symbol and the expression in black. In the summer of 2017, MathCheck was modified to print the relation symbol in magenta in those rare cases where there is strong evidence but no certainty of an error, or where the checking was less reliable than usual. Figure 2 shows an example of such a case.

$\sin x > -1$

No errors found. Warning! Some step looked bad, but MathCheck was unable to check it.

Figure 2. An example of when MathCheck has strong evidence of error, but not a certainty.

The example shows that MathCheck may fail to detect an error. Fortunately, addition, division, trigonometric functions, and so on have regular mathematical properties that make it unlikely for two different functions built from them to yield the same value for all the test values that MathCheck uses. Indeed, MathCheck has proven reliable in practice. Initially, the plan was to use a better proof engine like in STACK but testing with value combinations was needed in any case to produce counter-examples for the students, and when that had been implemented, it proved so reliable that there was no need to improve the proof engine.

In the equation mode, MathCheck checks that each step has at least the roots provided by the teacher, and after seeing the roots found by the student, MathCheck checks that they are also roots of the original equation. This makes it possible to deal with numerous equation types, instead of being restricted to, for instance, linear and quadratic equations. The array claim mode relies on checking with all arrays of size at most four with elements being integers between 0 and 3 (or between n and $n+3$, where n is an integer given by the teacher). In the propositional logic, quotient ring, and expression tree comparison modes, MathCheck checks the solution steps thoroughly. Also, membership of a string in the language defined by a context-free grammar is checked exhaustively. The comparison of context-free grammars given by the teacher and the student is based on generating strings in each language until a difference is found or an upper limit of work is met. It is thus incomplete.

MathCheck has been developed originally at the Tampere University of Technology (TUT) and then at the University of Jyväskylä (JYU). It has been open for student use since the autumn of 2015. Originally, MathCheck only had the simplification model illustrated above, without the graph-drawing feature that was added in December 2016 (Valmari, 2016, Valmari & Kaarakka, 2016). Since then, new problem modes have been added and old problem modes improved.

3 MathCheck as an area of research in the years 2015–2018

As the aim of the study is to improve MathCheck and confirm that it supports constructivism learning in university mathematics education, we have used design and development research (Richey & Klein, 2014) and design-based research (Anderson & Shattuck, 2012) methods. Our research is design-based as it contains an iteration process where interventions are used in traditional university mathematics education (Anderson & Shattuck, 2012). We have two types of interventions in our research: MathCheck itself as an educational tool and teaching modules, where MathCheck is used as a support. Within design and development research, MathCheck is a tool that is developed during the research and teaching modules are models that are studied during the tool development (Richey & Klein, 2014). The research was done in cycles in order to measure the learning outcomes of the students and to receive feedback on usability. Improvements have taken place in the form of new features and better instructions. The process is shown in Figure 3.

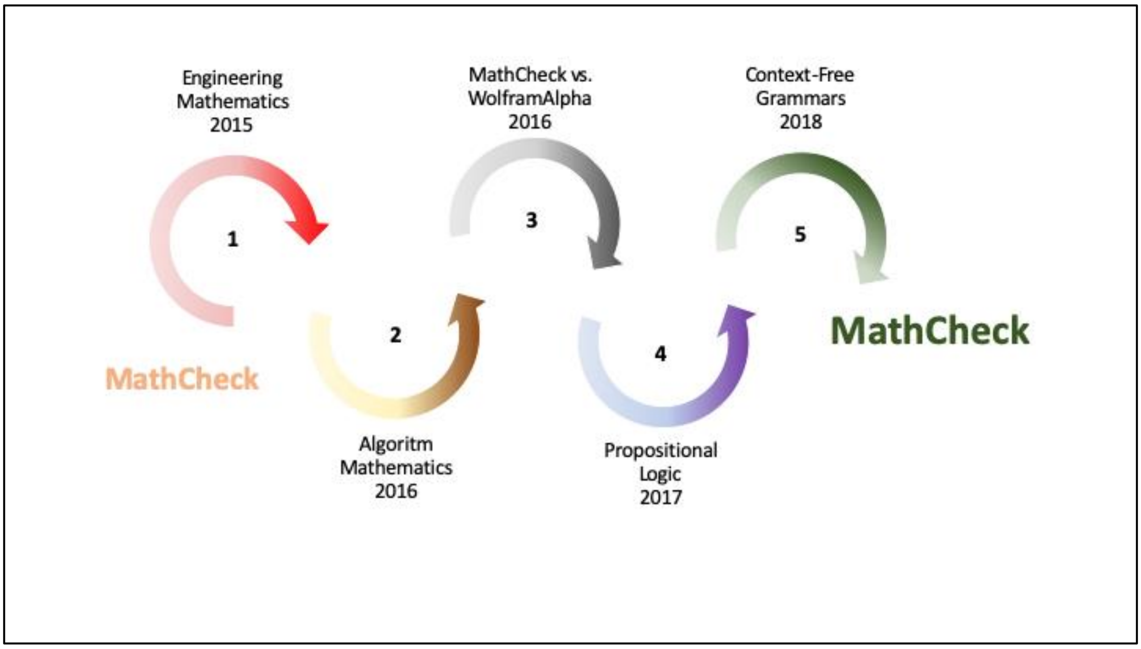


Figure 3. Research iterations used in this study.

The first iteration cycle, first experiment, was carried out in the autumn of 2015 and the latest in the autumn of 2018. Participation in the experiments has always been voluntary and experiments have taken place both in Finland and Norway.

The list below explains the interventions and aims of each intervention and connects them to the main research questions.

1. The experiment in Engineering Mathematics 2015 was carried out to receive information on the usefulness and user-friendliness of the first version of MathCheck. The feedback was used to improve the user interface.
2. The aim of the experiment in Algorithm Mathematics 2016 was to find out if MathCheck can be used to increase the understanding of expression approximation and time complexity. In general, the results would enlighten if MathCheck supports conceptual understanding.
3. The MathCheck vs. WolframAlpha experiment in 2016 compared student groups' learning outcomes when they used MathCheck and Wolfram Alpha. The results address both conceptual understanding and procedural fluency.
4. In the experiment in Propositional Logic 2017, the aim was to evaluate the usefulness of MathCheck as a supporting tool for independent studies in the basics of propositional logic and normal forms. The results gave answers to the second research question: "How can MathCheck empower both students and teachers in the education of mathematics?"
5. The Context-Free Grammars experiment in 2018 also addressed the "empowering of students and teacher" -research question as the aim of the experiment was to find out if MathCheck can be used to help a teacher to find a counter-example or be convinced that the CFG that is designed by a newcomer is correct.

3.1 Engineering Mathematics 1 in autumn 2015

Engineering Mathematics 1 was a first-year university-level course at Tampere University of Technology, TUT (Finland). Its contents included limits, continuity, and derivatives. In the experiment, students used MathCheck as a part of regular weekly exercises. Each week, one or two exercises among the full set of that week's exercises were MathCheck exercises, that is, their solutions were meant to be checked with MathCheck at home before the exercise session. The aim of using MathCheck was that students could check almost any solution and simplifications of intermediate steps with MathCheck on their own. The solutions were not returned to the teacher, that is, the use of MathCheck was solely between the student and MathCheck. However, the solutions were presented and discussed in the exercise sessions as usual. Figure 4 shows an example of a MathCheck exercise used in the course.

Simplify $\dfrac{(x-1)^2}{x^2-1} - 1$ assuming $|x| \neq 1$.

Figure 4. An example of a simplification exercise.

About 150 students used MathCheck in the exercises. Feedback on using MathCheck was obtained from 120 students. Of the students who gave feedback, 44 % experienced that MathCheck is useful, 40 % that it was not useful, and 16 % did not use it. The most common feeling was that MathCheck is useful.

3.2 Algorithm Mathematics in spring 2016

This experiment addressed an experienced teacher's (author) observations on first- and second-year university students' conceptual understanding in mathematics. The teacher had observed that approximating expressions from below or above is a tough task for the students making it also difficult to understand the concept of asymptotic time complexity (that is, the big O, Θ, and Ω notation). The following experiment was conducted to find out if MatchCheck can be used to increase the understanding of the expression approximation and time complexity. The participant group was the students in Algorithm Mathematics course at TUT.

Algorithm Mathematics is a first- or second-year course, depending on the student group. Its contents are set theory, relations, functions, logic, induction, and recursion. Because of the experience of the previous experiment, this time the students were given more difficult exercises to encourage the usage of intermediate steps in the solution process (Rasimus, Valmari & Kaarakka, 2016; Valmari & Kaarakka, 2016). These exercises were considered as special exercises instead of being part of the regular weekly exercises. The students were asked to save the feedback given by MathCheck as a PDF file and deliver it to the teacher via the course page in Moodle. For example, one of the problems was "Simplify the expression $f(x) = \frac{\ln((x^2+4x-12)^2)}{\ln(100)} - \frac{\ln(x+6)}{\ln(10)}$, and give the answer in terms of the log function."

In the same week with the exercise mentioned above, the students were asked to approximate the expression $\log(n^4 + n^3 - 5)$ upwards to find $c \in \mathbb{R}$ and $n_0 \in \mathbb{N}$ such that $\log(n^4 + n^3 - 5) \leq c \log n$ when $n \geq n_0$. Next week, the MathCheck exercises included the task of proving using the definition that (a) $2n^3 - n^2 + 5n = O(n^3)$ and (b) $2n^2 - 10n + 3 = \Omega(n^2)$. One question in the examination asked the students to prove that $\log(2n^3 - 6n^2) = \Omega(\log(n))$, using the definition.

Table 1 relates the points that students got from this examination question to the points that the same students got from the MathCheck exercises on $\log(n^4 + n^3 - 5)$, O and Ω. Each entry shows the number of students.

Table 1. Points on asymptotic notation problems.

	0 exam points	**1 exam point**	**2 exam points**
0 MathCheck points	49	8	14
1 MathCheck point	9	8	13
2 MathCheck points	4	8	22

The result shows that the MathCheck points that the students ($N = 135$) had obtained and the examination results had a positive correlation ($r = 0.4845$) which is statistically highly significant ($p < 0.001$). Unfortunately, this does not necessarily tell much about the benefit of using MathCheck. It is only natural that a skilful and motivated student performs better in both the MathCheck exercises and in the examination than a not so skilful and unmotivated student.

3.3 MathCheck versus Wolfram Alpha in autumn 2016

First-year students at TUT and the Norwegian Defence Cyber Academy (NDCA) participated in this experiment. The aim was to compare if students' learning with MathCheck and Wolfram Alpha differ. Both tools were used to check the correctness of solutions to simplification problems. Exercises and the final test are shown in Appendix A1-A3. Veera Hakala's (2016) project work contains more detailed results.

Altogether 146 students participated in the experiment, 106 in TUT and 40 in NDCA. In each place, the students were divided into two groups: those who were told to use MathCheck as a checking tool (N(TUT) = 56 and N(NDCA) = 20) and those who were told to use Wolfram Alpha (N(TUT) = 50 and N(NDCA) = 20). Each student had to solve a collection of exercises and check the solutions / final answers either with MathCheck or Wolfram Alpha. After completing the exercise collection, the students took part in a test, which was done without any tools. The maximum possible number of points from the test was 16. In the test, Finnish students were also asked to tell the time they had spent with the program. The students were divided into three grade intervals: 0 or 1, 2 or 3, and 4 or 5. The highest possible grade is 5, and the lowest accepted grade is 1. Table 2 and Table 3 show the grade distributions of MathCheck and Wolfram Alpha users in each usage time group among Finnish students.

Table 2. Finnish MathCheck users' grade distribution by time consumption.

Grade	t < 1h (*N*=26)	t ≥ 1h (*N*=30)
0 – 1	46%	20%
2 – 3	42%	50%
4 – 5	12%	30%

MathCheck users who had used the program at least one hour succeeded better than students who had used it less than one hour (Table 2). A similar difference cannot be observed among Wolfram Alpha users (Table 3). It can also be observed that among those who had used the tool at least one hour, 30 % of MathCheck users and 16 % of Wolfram Alpha users got one of the two highest grades 4 or 5.

Table 3. Finnish Wolfram Alpha users' grade distribution by time consumption.

Grade	t < 1h (*N*=19)	t ≥ 1h (*N*=31)
0 – 1	37%	39%
2 – 3	47%	45%
4 – 5	16%	16%

In Norway, all of the students (N = 40) used either Wolfram Alpha or MathCheck over an hour because they did their exercises during lessons. Therefore, Norwegian students belong to the category "used at least an hour". Half of the students used MathCheck and the other half Wolfram Alpha. Due to the small number of participants in Norway, it is not reasonable to analyze Norwegian results in isolation. In Table 4, the Finnish and Norwegian students' results have been combined.

Table 4. Grade distribution of MathCheck and Wolfram Alpha users in Finland (N = 61) and Norway (N = 40) who used at least an hour.

Grade	MathCheck (*N*=50)	Wolfram Alpha (*N*=51)
0 – 1	18%	24%
2 – 3	46%	53%
4 – 5	36%	24%

From Table 4 it can be seen that among the users of MathCheck, the proportion of students in the highest-grade interval 4–5 (36 %) is higher than the similar proportion with Wolfram Alpha (24 %). In brief, those students who practised with MathCheck succeeded better than those who practised with Wolfram Alpha.

The students at NDCA were asked an open question of whether MathCheck is a suitable tool for independent studying. Nineteen students out of twenty answered the questionnaire, and from those 14 thought that MathCheck applies well or to some extent for independent studying. Five out of nineteen students experienced that MathCheck does not apply for independent studying or is too hard to use.

3.4 Propositional logic (Algorithm Mathematics) in spring 2017

In the spring of 2017, MathCheck was experimented again in Algorithm Mathematics course with 160 participants at TUT. However, the focus was different. A teaching module was created containing the basics of propositional logic and normal forms. It was a part of the course, but the idea was that students could independently study and practice these topics with the module. Basics of propositional logic were familiar to the students from previous mathematics courses, so the propositional logic part was more of a revision. Conjunctive and disjunctive normal forms and full normal forms were new topics.

The module contained a total of eleven pages. The most common structure for a page was a short theory part, an example and an exercise about the current topic. This way the module was interactive and students got to try the theory immediately in practice. Figure 5 shows an example of the exercises of the module.

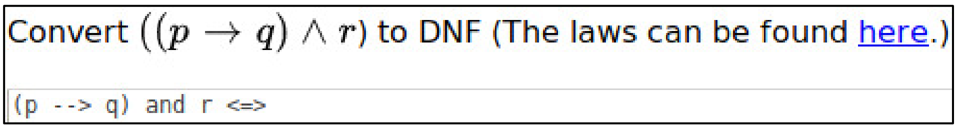


Figure 5. An example of Algorithm Mathematics exercise.

The students were motivated to complete the module by telling them that one examination question will be about one of the topics of the module. After completing the module, the students were asked the following six open questions:

1. Did you get all the exercises done in your opinion?
2. Did the module help your learning?
3. Was the platform pleasant to use (why / why not)?
4. Would you want to study independently with this kind of a platform in the future?
5. Would you have liked to have a pause option during the module?
6. Development proposals?

Unfortunately, only 21 students answered the questionnaire. The real number of who made the module cannot be known, because MathCheck does not keep any track of its users. According to the open questions, 20 out of 21 students announced that they had done all or almost all exercises. The number shows that the exercises have not been too hard and that those who have done the module have been motivated. From 21 students, 15 commented that the module had helped their learning, and only two said that it did not help at all.

The answers to the third question were categorized into three groups: positive (the platform was pleasant to use); positive but needs improvement, and negative. Nine out of 21 students experienced the platform as pleasant, six answered positively but felt that it could have been better with improvement, and five felt that the platform was not pleasant to use. One set of answers did not match the questions. The answers tell that the user interface could be improved.

Thirteen out of 21 students reported their willingness to study independently at home. Two out of 21 preferred that a part of the teaching would be independent. They preferred the blended learning method, where different kinds of teaching methods are used during the course. Two said that they could be interested in studying independently if improvements were made and two did not want to study with this kind of a platform. The answers of the two students did not match the questions. It seems that the majority of students who answered would like to study independently, but considering the small number of participants, it may be that the students who did the module, were already motivated to study independently.

The module did not have an explicit opportunity to pause, because implementing such a possibility would require introducing user accounts, which we want to avoid. Indeed, MathCheck does not collect any data about its users and does not know who is using it. The module was not very long, so it should have been possible to complete it during one session. Furthermore, one can save the position by using a feature that is available in all browsers: by bookmarking the current question page. Still, 17 out of 21 students would have wanted a pause option, which could also result from the fact that by mistake, students did not know the overall length of the module. A couple of students commented that this module was of suitable length, but any longer would have needed a pause option.

There were many development proposals. Despite the small number of answers, some problems came up often. In addition to the pause option, students suggested that the program should point out more precisely the location of the error and that the screen view should be more modern. Also, it was proposed that MathCheck should check not only that the answer is logically equivalent to the correct answer, but also that it satisfies the particular requirements stated by the teacher, such as if it should be in the disjunctive normal form.

3.5 Context-free grammars (Automata and Grammars) in autumn 2018

Context-free grammars (CFGs) are the most important method of defining structures of formal languages, such as programming languages. They are a simple but deep mathematical formalism. If a CFG does not yield the intended language, then there always is a counter-example. A CFG designed by a beginner is sometimes so difficult to analyse that the teacher can neither find a counter-example nor be convinced that the CFG is correct. This makes the teaching of CFGs difficult.

In the autumn of 2018, features were added to MathCheck for comparing the languages defined by two CFGs, checking whether a character string belongs to the language defined by a CFG, and for drawing a parse tree in case it does. A web page that teaches CFGs and contains exercises was written. Students of the Automata and Grammars course at the University of Jyväskylä, JYU (Finland), were given a link to this web page among their weekly homework problems. The CFG exercises constituted one-third of the problems of that week, while two-thirds were traditional paper and pencil tasks. The CFG exercises are given the link in Appendix B1.

At the beginning of the next meeting, the students were given a questionnaire in the form of a piece of paper and asked to fill it immediately. The questionnaire is shown in Appendix B1. Altogether 28 students returned it fully (25 students) or partially (3 students) filled. In every case, at least 14 out of a total of 18 questions were answered. Eighteen students told that they had done at least 80 % of the CFG exercises, 2 more had done at least 60 %, 6 more at least 40 %, and the last two at least 20 %. Table 5. shows the results for some of the questions.

Table 5. The results of the questionnaire. The columns are sd=strongly disagree, wd=weakly disagree, n=neutral, wa=weakly agree, sa=strongly agree, a=average, and p=statistical significance (p-value). The limits for *, **, and *** are 5 %, 1 % and 0.1 %, respectively.

Question	sd	wd	n	wa	sa	a	p
The exercises are suitable for 1st-year students	1	3	8	10	6	3.6	*
The exercises are suitable for 2nd- and 3rd-year students	0	0	5	12	11	4.2	***
The exercises are suitable for 4th-year and older students	0	0	7	13	8	4.0	***
It was more pleasant to study with this than with traditional exercises	0	0	1	16	11	4.4	***
I believe I learnt more than I would have with traditional exercises	0	0	1	18	9	4.3	***
The exercises make traditional lectures on the same topic unnecessary	7	15	6	0	0	2.0	***

The students had to collect a sufficient number of points from the weekly meetings to earn the right to participate in the examination. A student got points by telling in the meeting what exercises they had done and/or by actively participating in the discussion on a solution. Seven students claimed points only from the CFG exercises, three only from the remaining exercises, and 18 from both. That is, the students favoured the web-based exercises over the traditional exercises. At least three students who had returned the questionnaire did not claim points from the CFG exercises, perhaps because of doing too small a percentage of them, or because of not bothering (if they already had many enough points). Among the students who claimed points that week, 18 had and 10 had not already earned enough points meaning that the sample represents both fast and slow students.

It is clear that the students liked the MathCheck CFG exercises. Unfortunately, observations made later in the course and after the examination revealed that the students had not learnt the topic as deeply as the teacher hoped. Since then, much more MathCheck-based teaching material on CFGs has been developed. After all, exercises worth one-third of a week are not much for a topic like CFGs.

4 Discussion

Hundreds of university students have used MathCheck in their mathematics courses during the five experiments presented above. Generally, the feedback on using MathCheck collected via inquiries and interviews has been positive. This chapter discusses the results of the experiments in the light of the research questions.

4.1 How can the usage of MathCheck support the aspects of conceptual understanding and procedural fluency of constructivism learning?

The results in Experiment 2 (Algorithm Mathematics 2016) and Experiment 3 (MathCheck vs. WolframAlpha 2016) indicate that using MathCheck when evaluating own solutions helps students to gain conceptual understanding and increase procedural fluency.

The nature of using MathCheck differs from that of common mathematical programs that are used in teaching, for example, Wolfram Alpha or STACK. MathCheck is to be used during the solution process, for checking whether the intermediate steps are correct. The student can develop the solution step by step and check each step immediately (or rather the sequence of steps written so far). MathCheck points out errors but does not tell what the right step would be. So, the student must oneself analyse what the possible mistake is. As a consequence, MathCheck directs better towards conceptual understanding than Wolfram Alpha or even STACK.

MathCheck supports procedural fluency because when a student is given many exercises (whose solutions can be checked by the student herself), the student has to pay attention to writing expressions precisely and with several repetitions, the fluency will increase. In contrary, with Wolfram Alpha, the student has only to write correctly the starting point, and the program does the rest independently.

In more detail, MathCheck proved especially suitable when approximating values of functions upwards or downwards. Students are used to computing with precise values. However, in the real world, it is often necessary to approximate values rather than calculate with precise values.

With simplification problems, MathCheck seemed an excellent tool for students. The possibility of making mistakes increases with the number of computing steps. Similarly, finding the mistakes becomes more difficult the longer the path to the final solution is. With the help of MathCheck, the mistake is quickly found and the limited time can be spent on solving the problem instead of being wasted on finding the first error. Also, it was observed that when the students were forced to define the domain (e.g., declare $x \neq -5$ if the expression is $1/(x+5)$) before checking the answer with MathCheck, the habit stuck and several students continued to define the domains through the whole course. This is an improvement because usually, this habit fades away when it is no longer "needed" meaning that it is not noted in the book's solutions.

4.2 How can MathCheck empower both students and teacher in the education of mathematics?

Students mostly experienced MathCheck as a useful tool in mathematics education. However, not everyone found MathCheck as useful, especially not in the beginning. Not understanding the scope of MathCheck explains partly why a large number of students participating in the first experiment did not experience MathCheck as useful. Some students did not understand that the idea is not that MathCheck should find the final answers for them, but the idea is that MathCheck should give them feedback on their solutions. One factor may also have been too easy tasks. As it happened, some of the exercises used in the experiment were too simple, so there were no intermediate steps that needed checking (Rasimus & Valmari & Kaarakka, 2016; Valmari & Kaarakka, 2016).

In the rest of the experiments, the scope of MathCheck has been clearly explained and the complexity of the exercises has been raised.

As stated earlier in the MathCheck vs. WolframAlpha experiment in 2016, those who used MathCheck succeeded better in the examination than those who used Wolfram Alpha or no tool at all (where the dividing line between "used" and "not used" is one hour). The same, that is, usage of MathCheck improved examination results, was also noticed in other experiments (Algorithm Mathematics 2016 and 2017) when comparing the students' activity on doing MathCheck exercises. However, it has to be taken into account that also other factors such as motivation affect the examination results.

In Propositional Logic 2017, MathCheck was used as a supporting tool for self-study of the basics of propositional logic and normal forms. Most of the students who answered the questionnaire in the course commented that the independent learning module had helped their overall learning. Similarly, most of the respondents reported their willingness to study independently at home. However, in order to gain the full benefit of MathCheck in independent studying, thorough user guidance is needed to be given.

From the teachers' point of view, MathCheck decreases the teachers' workload, especially with courses of a large number of students. For example in exercise sessions, MathCheck, instead of the teacher, can show the exact point of the mistake. One suggestion for lowering teachers' workload was evaluated in the Context-Free Grammars experiment in 2018 where MathCheck was used to help a teacher to find a counter-example or to be convinced that the CFG that is designed by a beginner is correct. It became clear that the students liked the MathCheck CFG exercises; however, the number of homework problems was too small in order to gain a deep understanding of the topic as it was hoped.

MathCheck also offers an alternative for differentiating the level of education based on the students' individual abilities. Teachers can create extra problems for those students who need or want extra practice. It is possible to build web pages that create random problems of a fixed structure but varying parameters. By creating exercises of different levels of difficulty, MathCheck can be used as a differentiating method, thus taking the students into account, no matter what their starting level is. Another way the teachers can use MathCheck is to create teaching modules or courses. Teaching modules can be used as a revision or as a tool for learning a new topic. The modules give students more flexibility, in that they can decide when and where they will study.

4.3 User interface

The user interface issue deserves a discussion. Some of the first-year students had problems with textual input. We investigated the possibility of adding a mouse-clickable keyboard to question pages. It can be used for selecting the most commonly needed symbols and structures. For instance, when clicking √ it would write sqrt() into the answer box so that the user can write the argument between the (and). One problem is that such a keyboard can only contain a small number of symbols because otherwise, it would occupy too much space on the question page.

With Norwegian participants, there were fewer problems with textual input. There are two explaining facts; several Norwegian participants had earlier programming experience (because among selection criteria for entering to NDCA, programming experience is counted as positive) and that all Norwegian participants had a programming course on the same semester than the mathematics course where the experiment was conducted. It seemed that motivation for programming generally helped to adopt a new program with textual input.

The students at JYU studied information technology. They had no serious problems with textual input.

There also is another user interface issue. Technically, MathCheck is executed via web forms. It stores information neither on the server nor on the user's computer. There is no need for downloads or opening an account. Starting to use MathCheck as a student is technically as easy as it can be. Furthermore, question pages are just ordinary web pages with a web form. Therefore, teachers that are fluent with HTML and CSS have very great freedom in making them be whatever they want. The other side of the coin is that the possibilities to provide feedback by MathCheck in a natural and easy-to-use fashion are limited.

Initially, MathCheck provided its feedback as a separate web page that replaced the question web page on the user's screen. Getting back to the question page was possible using the back button of the web browser. As an attempt to improve user experience, since April 2017, many question pages have contained two submit buttons, one that delivers the feedback as was described above and one that opens it to a new tab (or a new browser window, if the browser has been configured to work so). Therefore, the students can choose whichever feels better.

There have also been attempts to make the feedback open into a separate box that is beside the answer box. That is otherwise very natural, but introduces the need for clumsy scrolling, if the answer is too wide or long. Because MathCheck aims at making it possible to ask students problems whose solutions need many, possibly complicated steps, long and wide feedbacks should be expected.

In July 2017, we decided to test feedback function with giving the students two submit buttons, one that opens feedback in an area to the right of the answer box, and another that opens it in a new tab or browser window. The idea was that the students always first use the former button, and then use the latter if scrolling becomes a problem. Submitting the same answer twice is not a problem and does not force to rewrite the answer.

This improvement made it possible to put many exercises on the same web page, together with text that teaches the material in question. Consequently, question pages grew long. Originally, many feedback boxes were used, each one beside the group of questions that it corresponds to, so that when the long web page is scrolled, always the relevant feedback box is visible. In January 2018, we found out how to fix the position of the feedback box, that is, it does not move when the question page is scrolled. The question pages written since then contain only one feedback box. Each group of questions has two submit buttons, one sending the feedback to the feedback box and the other sending it to a new tab or window. Figure 6 shows an example.

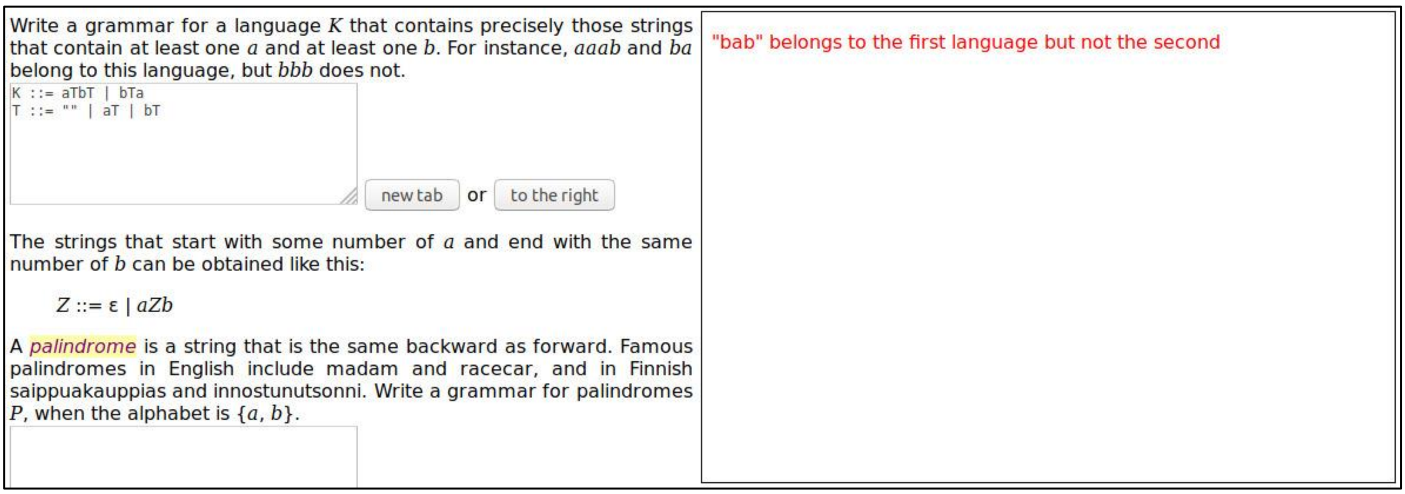


Figure 6. An example of the new user interface.

Since these improvements, the students have made very few complaints on the user interface. Unfortunately, now that it is possible to put multiple question groups on the same question page, a new problem arose: it is technically challenging to combine the answers to different question groups into a single package that could be sent to the teacher or a point recording system. With Firefox, it is possible to save the page in such a way that the resulting file contains all the answers (and also the questions, which is an advantage), but we have yet not found out how the same could be achieved with other browsers. Making it possible to save the answers one group at a time would be technically easy, but this solution is clumsy for the students. It may be that a reliable solution to this problem is only possible when using user accounts. One reason why we have not put much effort in solving this problem is that, for reasons explained by (Gibbs & Simpson, 2004; Gibbs, 2010), the authors believe that it is not necessarily pedagogically advantageous to record points in the middle of a course.

In the first four experiments, when asked about the user experience, students would have wanted more precise feedback from the location of mistakes. They also commented on a bit outdated screen view and hoped it to be updated to become a bit more modern and pleasant to the eyes. Besides, a feature that would check whether the answer satisfies all special requirements in the teacher's question was hoped. Most of these issues had been addressed by the fifth experiment. Consequently, similar remarks were almost absent in the feedback obtained from the fifth experiment and from other users of MathCheck by more than 100 students at JYU.

Students in JYU wanted answer boxes to have a running number so that it would be easier to refer to the right place when discussing an exercise. This has now been implemented. Currently, the only repeatedly occurring wish is that there should be a mechanism for recording the answers so that the students could more easily reproduce their answers in the weekly meetings of a course. Our standard reply is: With such a mechanism, you would run into trouble in the examination because the recorded answers would not be available there. Therefore, the idea is not to record the answers but to learn the topic so well that you can re-generate the answers.

The present version of MathCheck has seven problem modes: simplification of arithmetic expressions (including derivatives), propositional logic, equation solving, use of predicate logic for formulating claims about arrays, predicate logic and equations in quotient rings, expression tree comparison (a problem mode designed to help students to perceive expressions as structural entities and learn such concepts as operator precedence) and context-free grammar (also known as Backus-Naur form). In the experiments reported in this study, the simplification, propositional logic, and context-free grammar modes were used.

5 Conclusion

The results from above are only suggestive, but they are encouraging. Overall study shows evidence that MathCheck supports conceptual understanding and procedural fluency. The results also indicate that MathCheck can be used as a supporting tool in individual studies. In addition, MathCheck can lower the workload of a teacher.

The interest in MathCheck is growing in the Mathematics laboratory of TUT. In August 2017 MathCheck was connected to the electronic examination system Exam. During the examination, MathCheck only checks that the solution is syntactically correct and satisfies the particular requirements stated by the teacher, for instance, is in disjunctive normal form and is not more complicated than allowed. Afterwards, the teacher can use the full checking ability of MathCheck making the grading process quicker (and perhaps even more reliable) while reducing the teachers' workload.

In the future, the university education will highlight more student-oriented teaching, where the aim is constructivism learning facilitating a deeper conceptual understanding of mathematical concepts and procedures (Rämö, Oinonen & Vikberg, 2015). MathCheck addressed this and it has a role both as a part of traditional university courses (lectures, practice sessions) and as a supporter of the students' independent studying. No matter the place or time, students can use MathCheck during the solution process to check the correctness of the part of the solution obtained so far. The same applies to both teacher-given problems and problems that the students invent by themselves.

References

Anderson, T., & Shattuck, J. (2012). Design-based research. *Educational Researcher*, 41(1), 16-25. https://doi.org/10.3102/0013189X11428813

Bada, S.O. (2015). Constructivism learning theory: A paradigm for teaching and learning. *Journal of Research & Method in Education*, 5(6), 66-70. https://doi.org/10.9790/7388-05616670

Boudourides, M.A. (2003). Constructivism, education, science and technology. *Canadian Journal of Learning and Technology*, 29(3).

GeoGebra, (n.d). Retrieved April 1, 2019, from https://www.geogebra.org/

Gibbs, G. (2010). *Using assessment to support student learning*. Leeds Met Press.

Gibbs, G., & Simpson C. (2004). Conditions under which assessment supports student earning. *Learning and Teaching in Higher Education*, *1*, 3-31.

Hakala, V. (2016) MathCheck ja Wolfram Alpha opiskelun tukena (Unpublished report). Tampere University of Technology.

MathCheck (n.d.). Retrieved April 1, 2019, from http://users.jyu.fi/~ava/mathcheck.html

MathWorks Matlab (n.d). Retrieved June 1, 2019, from http://se.mathworks.com/products/matlab/

Maxima (n.d). Retrieved April 1, 2019, from http://maxima.sourceforge.net/

Mäkelä, A-M. (2016). Verkkotyökalut yliopistomatematiikan peruskursseilla (Master's thesis). Retrieved from TUT DPub of Tampere University of Technology. https://dspace.cc.tut./dpub/bitstream/handle/123456789/23768/Makela.pdf ?sequence=1

Mäkelä A-M., Ali-Löytty S., Humaloja J-P., Joutsenlahti J., Kauhanen J., & Kaarakka.T. (2016). STACK assignments in university mathematics education. In *Proceedings of 44th SEFI conference, engineering education on top of the world: industry university cooperation*. Retrieved from http://www.sefi.be/conference-2016/papers/Mathematics_and_Engineering_Education/makela- stack- assignments- in-university- mathematics- education- 73_a.pdf

National Research Council. (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: The National Academies Press. https://doi.org/10.17226/9822

Pelkola, T., Rasila, A., & Sangwin C. (2018). Investigating bloom's learning for mastery in mathematics with online assessment. *Informatics in Education*, 17(2), 363-380. https://doi.org/10.15388/infedu.2018.19

Rasimus E., Valmari, A., & Kaarakka T. (2016). Loppuraportti: Kaavat kuntoon! Kaavatarkastin opiskelijan apuvälineenä (Unpublished report). Tampere University of Technology.

Richey R.C., & Klein J.D. (2014). Design and Development Research. In J. Spector, M. Merrill, J. Elen, & M. Bishop (Eds), *Handbook of research on educational communications and technology*. New York, NY: Springer.

Rämö, J., Oinonen L., & Vikberg T. (2015). Extreme Apprenticeship – Emphasising conceptual understanding in undergraduate mathematics. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the ninth congress of the European society for research in mathematics education* (pp.2242-2248). Prague, Czech Republic.

Sangwin, C. (2015). Computer aided assessment of mathematics using STACK. In S. J. Cho (Ed.), *Selected regular lectures from the 12th international congress on mathematical education* (pp.695-713). Cham: Springer International Publishing. https://doi.org/10.1007/978- 3-319-17187- 6_39

Valmari, A. (2016). MathCheck relation chain checker. *In Proceedings of the Finnish mathematical days 2016* (pp.44–46). Retrieved from http://www.math.utu./materiaali/mathdays16/proceedings.pdf

Valmari, A., & Kaarakka, T. (2016). MathCheck: A tool for checking math solutions in detail. In *44th SEFI conference, engineering education on top of the world: Industry university cooperation*. Retrieved from http://www.se.be/conference2016/papers/Mathematics_and_Engineering_Education/valmari- mathcheck- 67_a.pdf

Valmari, A., & Rantala, J. (2019). Arithmetic, Logic, Syntax and MathCheck. In *Proceedings of the 11th international conference on computer supported education*. In H. Lane, S. Zvacek, & J. Uhomoibhi (Eds.), *Proceedings of the 11th international conference on computer supported education*, 2, (pp. 292-299). SciTePress. https://doi:10.5220/0007708902920299

Wolfram Alpha. (n.d.). Retrieved April 1, 2019, from https://www.wolframalpha.com

Weegar M. A., & Pacis D. (2012). A Comparison of Two Theories of Learning–Behaviorism and Constructivism as applied to Face-to-Face and Online Learning. In *Proceedings E-Leader Conference*. Manila.

Appendix A1: MathCheck vs. WolframAlpha 2016 - Exercises

You can use MathCheck whenever you have a mathematical formula that you are trying to solve. MathCheck has an input field, in which you can write your mathematical formula. There are also some demonstration and instruction links for further information. The link "How to type symbols and commands" includes lists of used commands and variables.

When using MathCheck you first write the formula chain into the input box on the left and then click the "Submit" button. MathCheck will then check the answer key and provide feedback. The colour of the feedback is dependent on the validity of your formula: - green, if your answer is logically right - black, if MathCheck cannot find any counterexample or proof. In this case your answer is most likely right. - red, if there are some mistakes in your answer. In case there are mistakes in your formula, the program will provide an example in which the answer doesn't hold true.

You should always solve the problem with pen and paper before using MathCheck. First you should write the formula conditions into the MathCheck input field. Write conditions between commands #assume and #enda. If there are many of conditions you can separate them using $\# \wedge \wedge$.

Example: #assume $x >= 08! = -2\#$ enda

If the error "Relation does not hold when..." occurs the reason might be a miscalculation or a lack of conditions.

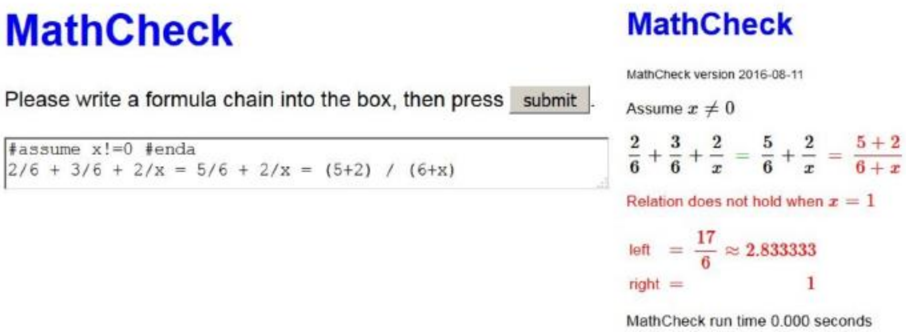


Figure 1: On the left you can see the MathCheck input field. On the right there is the output from MathCheck.

It is possible to give multiple input at the same time if you separate sentences using command #Newproblem. The command deletes conditions defined earlier for the following function.

You can find more instructions behind the link.

Exercises

You can mark the assignment finished once you have checked the results with MathCheck.

1. Simplify a function

   a) $f(x) = \frac{x^2+x}{x^2-1}$

b) $h(x) = \frac{x^2+2x-24}{x^2-16}$
c) $g(x) = \frac{x}{1-\frac{1}{1-x}}$

2. Simplify a function
   a) $f(x) = \frac{1}{x-1} \cdot (x - \frac{1}{x})$
   b) $h(x) = (1 - \frac{1}{x}) : (1 - x)$
   c) $g(x) = \frac{2x^3-4x}{x^2-4x}$
   d) $h(x) = (x^{a-1})^{a-1} \cdot (a^a)^{2-a}$

3. Simplify $x^{a-1})^{a-1} \cdot (x^a)^{2-a}$

4. Show $\ln(n^4 + n^3 - 5) < c \ln n$ to be true when $n > n_0$. Estimate realtion $\ln(n^4 + n^3 - 5)$ upwards using greater relation. After estimating report the values of c and n_0.

5. Show that $\log(\frac{a}{b}) = \log(a) - \log(b)$. Check proof using MathCheck and logarithm with base 10.

6. Simplify
   a) $g(x) = \log x + \log(x + 1) - \log(x^2 - 1)$
   b) $g(x) = e^{2 \ln x} - 2x^2$

7. Define composite functions
   a) $f \circ g$ and
   b) $f \circ f$ when $f8x) = x^2 - 1$ and $g(x) = \frac{1}{x^2+1}$

8. Define composite functions $f \circ g$ and $g \circ f$ and their range of definition. The functions are
   a) $f(x) = x^2 - 4$ and $g(x) = \sqrt{x + 4}$
   b) $f(x) = \frac{x^2}{9}$ and $g(x) = \sqrt{x\sqrt{x}}$

9. Define the inside function $g(x)$ of function $f \circ g$ and check simplification with MathCheck when $f(x) = \frac{2}{\sqrt{x}}$ and $f \circ g = 5x$, $x > 0$.

10. Define composite functions $f \circ g$ and $g \circ f$. Define the range of definition when $f(x) = e^{2x}$ and $g(x) = \ln x$.

Appendix A2: MathCheck vs. WolframAlpha 2016 – Exercises

You can find Wolfram Alpha behind this link. Write the problem you want to solve into the input field and press enter (or click equality button). Wolfram Alpha gives you information about the input: graph, values of roots, simplifications, solutions...

Before using Wolfram Alpha solve the problem with pen and paper. Then you can compare your results to the ones provided by Wolfram Aplha.

When simplifying a function, input the original function or expression to Wolfram Alpha. You should pay attention to the following sections in the feedback:

- "Input" displays the written function in symbolic form. Check that you have written it right.

- "Plots" shows graphs based on your input

- "Alternate forms" shows simplifications based on your input. Simplifications take accounts of all of the ranges of definition.

Compare your simplifications to Wolfram Alpha's.

Wolfram Alpha uses notation "*log*" for natural logarithm (ln). You should always provide the base information for the logarithm in order to avoid confusion. You can enter natural logarithm as "$\log_e$ and briggs logarithm as "$\log_{10}$.

When solving composite functions, you should enter the outside function as: $f(x) = x + 1$. Then you can tell Wolfram Alpha to solve the outside function $f(x)$ with the magnitude of inside function $g(x) = x^2 - 5$. This is achieved by writing: "$f(x) = x + 1$, solve $f(x^2 - 5)$" in the input field.

It is possible to write both outside function and inside function before solving composite function if you have updated Wolfram Alpha to Wolfram Alpha PRO (premium).

Exercises

You can mark the assignment finished once you have checked the results with Wolfram Alpha at:

1. Simplify functions
   a) $f(x) = \frac{x^2+x}{x^2-1}$
   b) $h(x) = \frac{x^2+2x-24}{x^2-16}$
   c) $g(x) = \frac{x}{1-\frac{1}{1-x}}$
2. Simplify functions
   a) $f(x) = \frac{1}{x-1} \cdot (x - \frac{1}{x})$
   b) $h(x) = (1 - \frac{1}{x}) : (1 - x)$
   c) $g(x) = \frac{2x^3-4x}{x^3-4x}$
   d) $h(x) = \frac{x^2-3x+2}{x^2-4}$

3. Simplify $(x^{a-1})^{a-1} \cdot (x^a)^{2-a}$

4. Show $\ln(n^4 + n^3 - 5) < c \ln n$ to be true when $n > n_0$. Estimate realtion $\ln(n^4 + n^3 - 5)$ upwards using greater relation. After estimating report the values of c and n_0.

5. Show that $\log(\frac{a}{b}) = \log(a) - \log(b)$. Check proof using MathCheck and logarithm with base 10.

6. Simplify

   a) $g(x) = \log x + \log(x+1) - \log(x^2 - 1)$

   b) $g(x) = e^{2 \ln x} - 2x^2$

7. Define composite functions

   a) $f \circ g$ and

   b) $f \circ f$ when $f(x) = x^2 - 1$ and $g(x) = \frac{1}{x^2+1}$

8. Define composite functions $f \circ g$ and $g \circ f$ and their range of definition. The functions are

   a) $f(x) = x^2 - 4$ and $g(x) = \sqrt{x+4}$

   b) $f(x) = \frac{x^2}{9}$ and $g(x) = \sqrt{x\sqrt{x}}$

9. Define the inside function $g(x)$ of function $f \circ g$ and check simplification with MathCheck when $f(x) = \frac{2}{\sqrt{x}}$ and $f \circ g = 5x$, $x > 0$.

10. Define composite functions $f \circ g$ and $g \circ f$. Define the range of definition when $f(x) = e^{2x}$ and $g(x) = \ln x$.

Appendix A3: MathCheck vs. WolframAlpha 2016 - Test

MathCheck green instruction/**Wolfram Alpha** yellow instruction

Name:
Student number:
High school (lukio) mathematics grade:
Graduate exam grade of mathematics (yo-koe):
What software did you practice:
How much time were you practicing:

Exercises

You can mark the assignment finished once you have checked the results with Wolfram Alpha at:

1. Simplify function $f(x) = \frac{x^3-9x}{3x^3-9x^2}$
2. Simplify function $h(x) = (1 - \frac{x}{1+x}) + \frac{1}{(}x^2 + x)$
3. Simplify function $h(x) = \frac{9-x^2}{x^2-6x+9}$
4. Simplify $\dfrac{a+\frac{b^2}{a}}{b+\frac{a^2}{b}}$
5. Simplify function $f(x) = \frac{\ln(x^2+6x)^2}{\ln 100} - \frac{\ln x+6}{\ln 10}$. And give the solution as a logarithm with base 10.
6. Simplify
   $f(x) = \log xy^2 - 2\log y$
7. Define composite function $f \circ g$, when $f(x) = \frac{x}{1+x}$ and $g(x) = e^x$, and calculate the value of composite function at point $x = 2\ln 2$.

Appendix B1: Context-Free Grammars 2018 – Exercises and questionnaire

29. Do half of the task on site http://users.jyu.fi/~ava/t_CFG.html.

30. Finish the task you started (see above 29. and site http://users.jyu.fi/~ava/t_CFG.html).

Context-Free Grammars 2018 – Questionnaire:

This questionnaire is a part of pedagogical education development project in October 2018. The questions are related to home tasks 29 and 30 in Context-Free Grammar course. The scale is as follows: 1=strongly disagree, 2=disagree, 3=neutral, 4= agree and 5= strongly agree. When percent is needed use scale: 1:$0 \leq x < 20$, 2:$20 \leq x < 40$, 3:$40 \leq x < 60$, 4: $60 \leq x < 80$ and 5: $80 \leq x \leq 100$.

Question:	1	2	3	4	5
1. There were too many difficult sub tasks.					
2. There were too many easy sub tasks.					
3. The whole task set was too difficult.					
4. The whole task set was too easy.					
5. The feedback received after incorrect answer was useful					
6. Praise and cheerful jokes after correct answers encouraged me to carry on.					
7. The exercises are suitable for first-year students.					
8. The exercises are suitable for second and third-year students.					
9. The exercises are suitable for fourth-year students.					
10. It was more pleasant to study with this than with traditional exercises.					
11. I believe I learnt more with these exercises than with traditional exercises.					
12. Going through the set of tasks in demostration class is useful (f.ex. then I will find out the correct solutions for the tasks I did not manage to do on my own)					
13. The exercises make the lectures on the same topic unnecessary.					
14. The exercises make the traditional demostration classes unnecessary.					
15. I used the drawing tool (found on right down corner).					
16. I could write context-free grammars before I did these tasks.					
17. I believe that I can now write context-free grammars.					
18. How many percents of sub tasks were you able to do?					

If you have comments, please write them down here.

Towards student-centred solutions and pedagogical innovations in science education through co-design approach within design-based research

Maija Aksela[1, 2]

[1] National LUMA Centre Finland, University of Helsinki, Finland
[2] The Unit of Chemistry Teacher Education, Department of Chemistry, University of Helsinki, Finland

The aim of this case study is to demonstrate how a co-design approach could be used within design-based research (DBR) with diverse multi-stakeholders in the LUMA[1] ecosystem to promote social creativity towards novel student-based solutions and pedagogical innovations. As a case, a national LUMA2020 development program (2019–2020), organized by the national LUMA Centre Finland and funded by the Finnish Ministry of Education and Culture, was studied in detail. The different data sources (e.g. an action plan, written observations) were analysed through qualitative content analysis. The Edelson's design-based research model used in the program offered a systematic framework or a map for co-designing both the action plan and its implementation. The co-design approach within the model was organised through three stages to engage all multi-stakeholders (altogether about three hundred participants) for it: (i) a research and societally oriented framework stage, (ii) a practical stage and (iii) a "bottom-up" stage in which teachers from 160 schools were active participants and professional key contributors. The co-design approach and the design decisions were facilitated by using guided face-to-face communication in small group work and digital creative learning spaces as a medium for social creative thinking. The co-designers were teachers, teacher educators, scientists or industry specialists in different stages. The co-design model used could be a way to bridge the newest research and innovations into praxis for supporting the curriculum at the school level and for promoting teachers' professional development by forming creative and diverse learning communities, in which all partners can learn from each other through sharing.

Keywords
co-design approach,
design-based research,
teachers' professional learning,
pedagogical innovations,
science education

Correspondence
maija.aksela@helsinki.fi

DOI
https://doi.org/10.31129/LUMAT.7.3.421

1 Introduction

"Together we are more!" (the LUMA[1] motto)

Design thinking is seen as central for promoting 21st-century competencies and practices in education (e.g. Noweski et al, 2012; Kelly et al, 2019). Enhancing social creativity (e.g. Fischer et al, 2005) and learning through a co-design approach with multi-stakeholders (e.g. teachers, students, scientists, teacher educators or industry specialists), could be a way to tackle multi-faceted challenges in science education and

[1] LUMA is abbreviated from "luonnontieteet, the Finnish word for natural sciences, and "matematiikka", the Finnish word for mathematics. The national LUMA* Centre Finland referred to here as the LUMA ecosystem with 11 universities (www.luma.fi) and about fifty partners.

its teacher training towards 21st-century competencies and student-centered solutions. Especially, creativity, collaboration and critical thinking are seen as necessary key competences.

There are many challenges in science education to be solved in the future. Science is not seen relevant enough for students themselves to study it at school or later on, especially in the developed countries (e.g. OECD, 2015). Their attitudes and interest have a big influence on their science enrolment behavior (e.g. Krapp & Brenzel, 2011; Regan, 2015). Relevant vocational and societal perspectives of science are often unknown. Although Finnish youth have been one of the most skilful students in science globally, their interest to study science is often very low according to the PISA results (Finland and PISA, 2019). School science should be promoted more positively for all – “perhaps as a ‘springboard' to new sources of interest and enjoyment.” (OECD, 2015, 6). More scientific literacy for all is also needed in the future, for example, to solve global challenges (e.g. climate, energy, food and water).

In addition, the 21st-century learning demands have to be better taken into account in the design continuum of science teacher education. How to strengthen teachers’ high professional role and teachers’ life-long learning (e.g. Niemi & Iso-Pahkala-Boureat, 2015)? How to get teachers opportunities to update their knowledge and skills concerning new research results in both science and its learning, thus to promote evidence-based teacher education for life-long learning in science (e.g. Aksela, 2010)? There is a need to bridge the gap between research and praxis (e.g. Juuti & Lavonen, 2006; Aksela, 2010; Anderson & Shattuck, 2012; Taber, 2017). Novel solutions for it are needed. Could the co-design approach within the design-based research (DBR) be a way to promote teachers’ life-long learning?

Teachers are seen as key professional contributors to reforms (e.g. Roschelle & Penuel, 2006). In Finland, teachers are valued and trusted as professionals in curriculum development, teaching and assessment (e.g. Niemi, Lavonen, Kallioniemi & Toom, 2018). They also have a lot of professional freedom to decide how to teach and collaborate within curricula. According to Juuti et al (2017), successful teachers’ professional development should be teacher-led, continuous (long-term), situated or connected to the classroom context, collaborative, and should include reflective practices. Design-based research (DBR) used as a design framework in the LUMA2020 program (see Section 2 for more details) has been earlier found as a useful way to promote teachers’ or future teachers’ professional development and growth (e.g. Sherin, 1998; Kelly, Lesh & Baek, 2008; Pernaa & Aksela, 2013;

Vesterinen & Aksela, 2013; Aksela & Vihma, 2015; Aksela et al, 2016; Juuti, Lavonen & Meisalo, 2016).

Facilitating the school-university partnership can be potential in contributing to the creation and translation of knowledge about teaching and learning (e.g. Baumfield & Butterworth, 2007), as it is the main aims of the LUMA* ecosystem. It could be especially useful to engage teachers in a long-term collaborative research agenda (e.g. Reeves, 2000). Teachers often fail to adopt pedagogical innovations, if they are designed only by researchers (e.g. Talbert & McLaughlin, 1999; Linn, 2006; Juuti & Lavonen, 2006). They mainly make decisions on their teaching based on their own needs (e.g. Zhao et al. 2002). It may have a positive effect on student achievement if teachers have a more active role in the co-design processes. Promoting the knowledge production of teachers points out: (i) shared an understanding of the challenge, (ii) a willingness to change one's own perspective, (iii) a commitment to participate in the dynamics of the group (Orland-Barak & Tillema, 2006).

The co-design approach focus in this study has led to high-quality teacher professional development for 21st-century learning used in a curriculum planning model (Kelly et al, 2019). Teachers can act successfully as co-designers with researchers (Roschelle & Penuel, 2006). How to facilitate the co-design approach and social creativity within diverse multi-stakeholders (e.g. teachers, teacher educators, scientists or industry specialists) towards novel solutions and pedagogical innovations in science education as in the LUMA2020 program? There is a need for more understanding of the co-design approach (see Section 3 for more details) for it to be successful. The aim of this case study is to understand the co-design approach within Edelson's design-based research model in the LUMA* ecosystem (see Section 4). Its research policy points out that the purpose of design-based research (DBR) is to create student-centred solutions with diverse partners (e.g. schools, industry) and share them in all school levels (Research and development policy of the LUMA Centre Finland, 2018). This case study focuses on the following guiding questions: (i) *how to facilitate the co-design approach?*, (ii) *who are the co-designers?*, (iii) *how can design decisions be executed in the process?* And (iv) *how to use the co-design process as a tool for promoting teachers' professional development?*

2 Design-based research as a framework for the co-design approach

Design-based research (DBR) has been found to be useful for developing new solutions and pedagogical innovations in education at least since the 1990s. By using it, educational practices are renewed through systematic, flexible and iterative analysis of design and development, and novel solutions are often produced for very complex challenges in authentic learning environments (e.g. Wang & Hannafin, 2004¸ Van der Akker, Kelly, Lesh & Baek, 2008). The term *design-based research* (e.g. Kelly, 2003; Juuti & Lavonen, 2006; Anderson & Shattuck, 2012) used in this paper has also been referred to in literature as (i) *design experiments* (e.g. Brown, 1992; Collins, 1992), (ii) *design research* (e.g. Cobb, 2001; Edelson, 2002), (iii) *development research* (e.g. Richey & Nelson, 1996), or (iv) *educational design research* (e.g. Sandoval & Bell, 2004; Van der Akker, Kelly, Lesh & Baek, 2008, Vesterinen & Aksela, 2013; Sandoval, 2014). Many kinds of successful models with various stages in practice have been reported (e.g. Lavonen & Meisalo, 2002; Clements & Battista, 2000). Usually, the design-based research has 7 to 9 different stages.

Design-based research usually gives us three kinds of information as a result of the study (Edelson, 2002): (i) information on the design product itself, (ii) the development process and (iii) the background theory or theories used in the development process. According to Edelson (2002) design methodology as a general design procedure provides guidelines for the process and describes (a) a process for achieving a class of designs, (b) the forms of expertise required, and (c) the roles to be played by the individuals representing those forms of expertise. As a result, concrete design solutions can be acquired: activities, materials, courses, learning environments, software or equipment for different levels (e.g. Brown & Campione, 1994; Cognition & Technology Group at Vanderbilt, 1997; Kelly, 2003). Some examples of design products are mentioned in the context of the LUMA ecosystem in Section 4.

Design-based research differs from traditional education research on the following eight areas: according to (i) *the role of the participants* (it involves different participants in the design to bring their differing expertise into producing and analyzing the design),(ii) *the amount of social interaction* (frequently it involves complex social interactions with participants sharing ideas, distracting each other etc.), (iii) *flexibility of the process* (it involves flexible design revision in which a tentative initial set is revised, depending on its success in practice), (iv) *characterizing the findings (*it involves looking at multiple aspects of the design and developing a profile that characterizes the design in practice), (v) *location of research conducted* (it often occurs in the buzzing, blooming confusion of real-life settings where most learning actually occurs), (vi) *the complexity of the variables* (it involves multiple dependent variables, including climate variables, outcome variables and system variables), (vii) *unfolding of procedures* (it involves flexible design revision in which a tentative initial set is revised, depending on its success in practice and (viii) *the object of research* (it focuses on characterizing the situation in all its complexity, much of which is not now *a priori*). (e.g. Barab & Scquire, 2004; Collins, 1999; Aksela, 2005)

The following characteristics of good design-based research guide its design and implementation process (Dede, 2004; Design-Based Research Collective, 2003): (i) the correspondence of the design in the needs of practical and education policy,(ii) the intertwining of the aims of the chosen intervention and developed theories, (iii) the cyclicity of the development between design, implementation, analysis and re-design, (iv) the reliability of received results, (v) how the outcome of the development works in an authentic environment and (vi) how the received results adapt to earlier theories and practical implementations. The validity of design-based research is shown often through collaboration (e.g. the results checked by other co-examiner(s) as in this case study) and iteration, and the reliability through using various references for the research and by evaluating the usefulness of the research concerning education and learning (e.g. Design-Based Research Collective, 2003; Edelson, 2002).

Design-based research can include a strong collaborative approach with various partners – the so-called co-design approach (see Section 3 for more details) in this paper. It is used here within the design-based research framework, called the Edelson model (Edelson, 2002; see Section 4 for details).

3 Co-design approach within design-based research

The co-design approach has been used since the early 1960s (Zamenopolous & Alexiou, 2018). During the years it has been applied to various fields, for example from computer software design to architecture. The co-design approach is close to many other traditions of design, for example, participatory design (e.g. Ehn, 1992; Lee 2008), learner-centered design (e.g. Soloway et al., 1994) or co-creation (e.g. Prahalad & Ramaswam, 2004). According to Zamenopolous & Alexiou (2018), co-designers can have different roles in the process: they can facilitate or engage others in design tasks or share, collect, interpret or create knowledge, ideas and resources, and also engage at different stages of a design project. Different kinds of technology (e.g. Living labs) can be used for facilitating co-design and implementing activities (e.g. Andersen, Kanstrup & Yndigegn, 2018).

The co-design approach has been found to be the most effective way to engage teachers in designing new practices at the school level (e.g. Penuel et al, 2007). Roschelle & Penuel (2006,1) define its use in education as "*a highly-facilitated*", team-based process in which teachers, researchers, and developers work together in defined roles to design an educational innovation, realize the design in one or more prototypes, and evaluate each prototype's significance for addressing a concrete educational need". The co-design approach can be seen as social (collective) creativity applied across the entire span of a design process (Sanders & Stappers, 2008). Co-designers can be, for example, scientists, teacher educators, teachers, specialists from industry.

The co-design approach has been found useful at the school level. It provides an opportunity to match the curriculum goals of teachers (Tissenbaum et al, 2012; Kelly et al, 2019) and increase reflections and ownership by a teacher (Roschelle & Penuel, 2006). Seven characteristic features are recommended to be taken into account when using co-design as an approach (Roschelle & Penuel, 2006): (i) it takes on *a concrete, tangible innovation challenge*, (ii) the process begins by *taking stock of current practice* and classroom contexts, (iii) it has *a flexible target*, (iv) it needs *a bootstrapping event* or process to catalyze the team's work, (v) it is timed *to fit the school cycle*, (vi) strong facilitation with *well-defined roles* is a hallmark of it, and (vii) there is central accountability for *the quality of the products* of co-design.

Although co-design with the DBR framework has been found to be very useful in producing various relevant solutions and also theories, as mentioned in Section 2, challenges may occur when using it. According to Piirainen, Kolfschoten & Lukosch (2009), five main challenges of collaboration can be found: creating shared understanding, balancing requirements of different stakeholders, balancing rigor and relevance in the process, organizing the collaboration effectively and creating ownership. If co-designers are novices, more guidance is needed to be successful (e.g. Chao, Saj & Hamilton, 2010). In addition, the following things could be taken into account: (i) the process is often time-intensive (e.g. Rheinfrank et al.,1992; Roschelle & Penuel, 2006), (ii) trust is needed on each other's knowledge and skills between co-designers (e.g. Shrader et al., 2001), (iii) criteria for success is needed between co-designers (e.g. Blomberg & Henderson, 1990), (iv) understanding of goals, roles, and contributions of each participant (e.g. Shrader et al., 2001; Lee, 2008), (v) tight integration of curriculum (e.g. Roschelle & Penuel, 2006) and (vi) understanding of negotiating shared frames during early design phases (e.g. Hey, Joyce & Backman, 2007). Designers' frames seem to be effective on design decisions and the actions that they will take (e.g. Schön, 1983).

4 Design-based research in the LUMA ecosystem

The aim of the national LUMA* Centre Finland (network of 11 universities and 13 LUMA Centres with around 50 partners; referred to here as the LUMA ecosystem) is to develop novel, student-centred, research-based solutions and pedagogical innovations, and to distribute them both directly and indirectly to all science education and learning on different educational levels (Research and development policy of the LUMA Centre Finland, 2018). The co-design approach is seen as central for its design-based research (DBR) framework. The first LUMA Centre was built in the year 2003 in order to build a bridge for promoting collaboration between universities, schools and industry (e.g. Aksela, 2015).

Design-based research has been used broadly in promoting science education or its research-based science teacher education earlier in Finland (e.g. Lavonen & Meisalo, 2002; Aksela, 2005; Juuti, 2005; Juuti & Lavonen, 2006; Pernaa, 2013; Vesterinen & Aksela, 2013; Juuti, Lavonen & Meisalo, 2016; Juuti & Lavonen, 2017). For example, relevant inquiry-based working instructions for science education have been designed collaboratively with diverse partners outside the university (e.g.

Aksela, 2005; Aksela & Boström, 2012; Aksela & Ikävalko, 2016). A Finnish book on design-based research contains other examples for carrying out design-based research in education and in science teacher education in Finland (Pernaa, 2013).

Design-based research is mostly connected to the studies and theses in science teacher education in the LUMA ecosystem. In research concerning doctoral theses, for example, the following new solutions and pedagogical innovations are produced with the help of design-based research (SECO, 2019): (i) learning games and a framework for their evaluation, (ii) inquiry-based working instructions in collaboration with future teachers and the industry, (iii) a science club model for small children's inquiry-based education, (iv) a model for teachers' educational development by using inquiry-based learning and SOLO-taxonomy, (v) a collaborative and engaging model for teacher education that promotes inquiry-based education in-class teacher education, (vi) a course in the context of the Nature of Science for future teachers, (vii) problem-based and inquiry-based laboratory work activities into university education, and (viii) molecular modelling activities for instruction.

In practice, design-based research (DBR) can be carried out in various ways (see Section 2), and different models are available for supporting design decisions carried out during design-based research (e.g. Sandoval, 2014). In the LUMA2020 program, the design-based research framework, the so-called Edelson's model (Edelson, 2002) has been applied in practice. It has two main parts that guide the process and the decisions of the process: (a) theoretical problem analysis and (b) empirical problem analysis (see Figure 1). In the different parts of its cyclic development process, the so-called mixed methodology is often used in order to understand the object of development and its relevancy based on design decisions. For example, video-recordings, naturalistic observations, group interviews, concept maps, learning diaries, students' research reports or surveys can be used (e.g. Aksela, 2005) through the co-design approach, especially with teachers (teachers as reflectors or researchers) in the framework.

The co-design approach of Edelson's model (Edelson, 2002) can be carried out systematically in the following steps within the LUMA ecosystem. The framework is also used in the LUMA2020 program (Figure 1; the main phases are marked bold in the text): (i) *mapping out the needs* for the development process together with the participants (often called **empirical problem analysis** or a needs analysis: it can be done through a survey with teachers or a content analysis of learning materials or curriculum framework; the needs can be national needs and/or teachers' needs in science education), (ii) *mapping out new research information* concerning the chosen theme and synthesis (**theoretical problem analysis**), (iii) *setting the goals of development* together with different stakeholders based on the steps i) – ii) (**goals for the activity**), (iv) *designing a pilot model* together (e.g. practical activities) for the object of development based on chosen aims, and *testing the pilot model* with the target groups and refining it based on received results (**a pilot model and testing it; an iterative design cycle**)

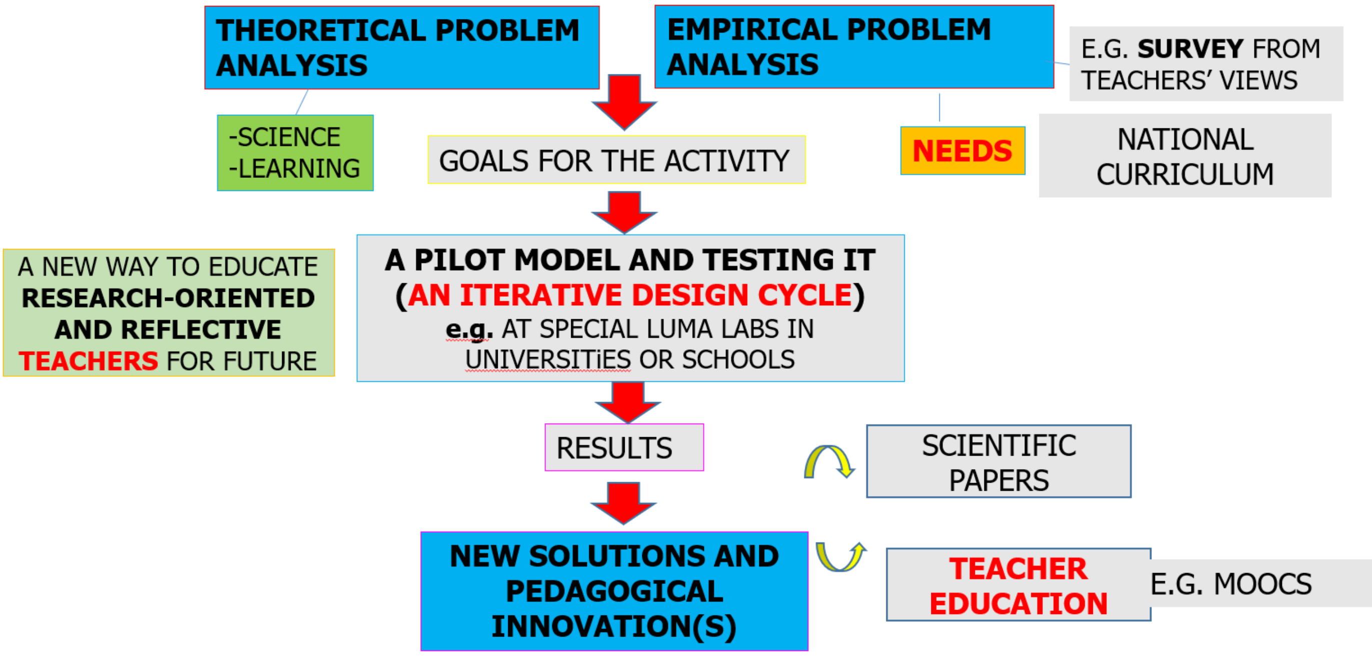


Figure 1. An example of the different phases of co-design within design-based research carried out in the Finnish LUMA ecosystem (see www.luma.fi) by applying Edelson's (2002) model. The different phases are marked bold in the text.'

(a cyclic model; teacher as a researcher or a reflector), (v) *describing the outcome* of development, and *reporting* it (**results and pedagogical innovations**) and (vii) *spreading new openings and solutions*, and *offering education* (e.g. through massive open online platforms, the so-called MOOCs) on them (**teacher education; scientific papers**). Usually, a researcher at a university, a teacher educator or a future teacher acts as a facilitator that carries out the synthesis and maps out new research information concerning the topic for other partners of the program or projects. In co-design meetings, steps (i) and (ii) are gone over together, and the aims for development and the model for implementation with timetables are arranged together.

Besides formal learning environments at schools, non-formal learning environments (e.g. 15 LUMA labs), are often used during the co-design processes, include science and technology activities for children, youth and entire families, such as clubs, camps, parties and events, as well as the pursuit of hobbies at home. For example, in ChemistryLab Gadolin (one of the LUMA Labs), new openings in the contexts of everyday chemistry, sustainable chemistry, and development and modern technology are developed together with visiting school groups (Aksela et al, 2018) within industry collaboration (Aksela & Ikävalko, 2016).

The distribution channels of the LUMA ecosystem include the education of future and current teachers at universities, events organized by universities and other partners, academic and popular multimedia publications, as well as international researcher exchange and education export. Innovations are spread to be used in non-formal, in-formal or formal learning environments. Research results will be published for the academic community in the form of articles in domestic and international peer-reviewed open access publications, conference presentations and proceedings, as well as scholarly works (bachelor's, master's and licentiate theses, doctoral dissertations).

The LUMA ecosystem has also channels of its own, such as the national LUMA days for teachers, International LUMAT Symposium and the peer-reviewed LUMAT (International Journal on Math, Science and Technology Education) online journal and the LUMAT-B online journal focused on conference and project proceedings, as well as the LUMA News section of the LUMA website. (Research and development policy of the LUMA Centre Finland, 2018). These acquired solutions are spread into teaching through teachers' pre-service and in-service education. As future teachers and teachers at schools have participated in designing, implementing and reflecting on the results of the development process, this acts as a novel model for organizing teacher education. An online book (Aksela, Oikkonen & Halonen, 2018) gives a summary of examples of the projects that have been carried out at the University of Helsinki since the year 2003.

5 A case study in the context of the LUMA ecosystem

The aim of this case study is to demonstrate how a co-design approach could be used within the design-based research (DBR), the Edelson's model (explained earlier in Section 4) with diverse multi-stakeholders (altogether about three hundred participants) in the LUMA* ecosystem to promote social creativity towards novel student-based solutions and pedagogical innovations focusing on the following guiding questions: (i) how to facilitate the co-design approach?, (ii) who are the co-designers?, (iii) how are design decisions in the process executed? In addition, the aim is to demonstrate (iv) how to use the co-design process as a tool for promoting teachers' professional development.

As a case, a national LUMA2020 development program (2019-2020) organized by national LUMA Centre Finland and funded by the Finnish Ministry of Education and Culture was studied in detail. The quality of the program has been guaranteed by using the best specialists in the evaluation process during the program and applying their advice on the design processes. The main aim, target groups and the design products of the LUMA2020 development program are summarised in Table 1. The main principles are given for the design process and the partners of the program can be found in Table 3 (Appendix 1). The organization of the program and responsibilities of different partners, and the stages of the program process can be found in Table 4 (Appendix 2).

The program is a continuum for the earlier national LUMA Suomi development program (2013-2019; www.luma.fi) funded by the Ministry of Education and Culture. The main focus of the program was the lower secondary level (6 to 16 years-olds). The focus of the LUMA2020 is especially on early childhood education and the upper secondary level, also in vocational education and training. The LUMA2020 program was chosen into this study because the co-design process framework has been documented in detail, and thus it is suitable for the content analysis method used (see later in detail).

Table 1. The main aim, target groups and the design products of the LUMA2020 development program (the text is translated from the action plan written in the co-design stage 1 by the LUMA ecosystem).

The main aims of the LUMA program (given from the policymakers)	**Target groups** (given from the policymakers)	**The design products** in the action plan designed by the LUMA ecosystem
As the program ends at the end of 2020 the program has: **-increased fascination** towards studying LUMA subjects and has improved the **quality** of teaching and learning from early childhood education to universities **-increased children and youths' interest** in LUMA subjects and their study and career possibilities (individual, vocational and societal relevance) **- strengthened the contents of teaching and learning, and methods in teacher training** at faculties of science in universities (early childhood education, class teachers, special needs education, subject teachers, teacher education that universities offer, vocational teacher education) - **promoted development work between faculties** of science / technical faculties (according to the subject), teacher training institutions and teacher training schools, and university of applied sciences and vocational teacher education -increased **competences of staff** and their own education in these institutes **To develop**: 1) children and youth's **formal** learning from early childhood education to secondary education. 2) **children's, youth's and families' free-time non-formal/informal science and technology education** and 3) **the competences of educational/ teaching staff**	The target groups of the actions include **3-19-year-old children and youth** – both girls and boys, **their guardians** and **educational/teaching staff** working on different levels from early childhood education to universities. The program includes the development of LUMA subjects' teaching and learning from early childhood education to universities, **by stressing** actions especially from **early childhood education to the upper secondary level** (also in vocational education and training).	New **operating models** are developed in the program for e.g. collaboration between early childhood education / upper secondary school / vocational institution and universities, working life collaboration, and **online courses**. **Virtual clubs** (packages) and **online courses for science and technology education** (national, shared between universities). These are exploited in order to strengthen the continuity/path of science and technology education from the very young to the very old. In addition, alongside the implementation of the program, **theses and other research papers** are written. **The final report of the program** that **includes evaluation** will be finished at the end of 2020. -building a new national **network of LUMA development communities** (e.g. with 50 partners in the LUMA ecosystem) -strengthening the **national LUMA contact person's network of municipalities.**

The different data sources (an action plan as a main source (see Table 3 and Table 4 in Appendix 1 and Appendix 2), a memorandum of the co-design meeting in a wiki platform at the University of Helsinki, the materials in the open web page (http://2020.luma.fi) and written observations by a researcher) were analysed through qualitative content analysis to understand the co-design processes within the model. Applying content analysis from the texts (Huberman & Miles, 1994), the central features of the co-design process and providing answers for the above-mentioned questions have been presented as the results in Section 6.

The co-design approach was facilitated in three stages (the answer for the first research question): (i) *a Research and societally oriented framework stage* (see Section 6.1)*, (ii) a practical stage* (see Section 6.2) and (iii) a *"bottom-up" stage* (see Section 6.3). As an example of the analysis, the form of the names of the stages are described (see Table 2)*:*

Table 2. An example of the content analysis.

The source 1	**The source 2**	**The name of the stage** formed from source 1 and source 2
The action plan: Table 3: LUMA Centre Finland (**a network of 11 universities**), In the implementation of the program, LUMA Centre Finland carries out development, education and marketing/communication collaboration with e.g. partners that are represented in the national LUMA advisory board. *=> **the role of the universities** is to bring the newest research to the co-design approach.* *=> "a research oriented"*	The LUMA2020 webpage: The names and organizations of LUMA advisory board -about 50 partners outside of university (e.g. industry) =>*"societally oriented"*	*a research and societally oriented framework stage*
A memorandum of the meeting (saved in the wiki platform): You can attend of the Facebook group to discuss more about the program: www.facebook.com/groups/LUMA2020. *=>the practical decisions of the action plan* *=> "**a practical stage**"*	The written observations: Few suggestions for a digital platform (e.g. Wiki, Teams and Facebook). *=> the practical decisions of the action plan* *=> "a practical stage"*	*a practical stage*
A memorandum of the meeting (saved in the wiki platform): The meetings and discussions with teachers will be organised once a month… =>the discussions of the program with teachers =>*a **"bottom-up" stage***	The written observations: The next teachers at each school are written a plan in the context of their school curricula in a digital form during a month. The co-designing will continue with the LUMA workers and other teachers from different schools during the next meeting." =>the teachers in each school made their plan for the program in details =>*a "bottom-up" stage*	*a "bottom-up" stage*

The names for the main phases of the co-design approach by the Edelson model framework (empirical **problem analysis, theoretical problem analysis**) has been formed from the text in Table 3: "*1) mapping out needs together with participants (empirical problem analysis, the so-called needs-analysis), 2) mapping out new research information concerning the chosen topic from sciences, their learning and teaching (theoretical problem analysis).*" In addition, the name for the third main phase, a cyclic development process was named from the stages 3-6: "3) setting the aims for development together with the participants based on steps 1 and 2, 4) planning a pilot model for the object of development (e.g. an activity, material) based on set aims, 5) testing the pilot model with the target group and refining the model based on received results (multiple steps), describing the development output and reporting and 6) spreading out new openings and solutions and offering education for these new topics. During the LUMA 2020 program, the development process is carried out at least in one cycle."

The written observations by a researcher from the stages (see Section 6.1, 6.2 and 6.3) were used to open more the texts in the main sources, for example, "*The co-designers suggested few suggestions for a digital platform (e.g. Wiki, Teams, Facebook).*" There it was mentioned only generally "*digital platform*" in the texts.

Because a researcher of this case study has been actively involved in the LUMA2020 program, a co-examiner has checked and accepted the written texts in this paper in order to increase the validity and the reliability of the case study.

6 Results and discussion

The co-design approach used within the design-based research framework, the Edelson's model (Edelson, 2002) is described in the following Sections (see Sections 6.1, 6.2 and 6.3) by providing answers to the following questions: (i) how to facilitate the co-design approach?, (ii) who are the co-designers?, (iii) how are design decisions executed in the process? The results for the question (iv) how to use the co-design process as a tool for promoting teachers' professional development? is presented in Section 6.4.

6.1 A Research and societally oriented framework stage

The characteristics of good design-based research have guided the design and implementation process of the LUMA2020 program, as described in Section 2. The co-design approach was used in the designing of the general action plan for the framework given by the policymakers through the Edelson model's three main phases (see Figure 1):(i) *empirical problem analysis* (the needs for co-design), (ii) *theoretical problem analysis* (most novel research in science and science learning) and (iii) *cyclic development process* (actions for the decided goals). Co-designers in stage 1 were teacher educators and researchers from 11 universities (13 LUMA centres) by facilitating a national team (a director and a coordinator of the centre, a chair and a vice-chair of the board). The digital platform *google docs* was used for writing the action plan with different stakeholders around Finland. First, the director and the coordinator of the program wrote the framework of the action plan and then other members of the LUMA ecosystem continued the writing process.

The design decisions were accepted first by the board of LUMA Centre Finland (a member of each 13 LUMA Centres) and then by the steering group of the policymakers (including invited members from the universities: a director of the program, a project manager and four special experts). Design decisions were made based on the co-designers' expertise (e.g. most novel research in the field), international assessment programs (e.g. TALIS, TIMMS, PISA), the new national curriculum framework, experiences of the earlier national LUMA Suomi program (the program continuum for the earlier one) and the ideas collected through brainstorming from about 50 LUMA steering group members (e.g. industry foundations and pedagogical teacher organizations).

Four themes for the program were chosen through the co-design approach (i) sustainable development (e.g. climate change), (ii) math around us (e.g. math and art), (iii) technology around us (e.g. Al) and my LUMA (open for different integrated topics over subjects). A successful international StarT program (see https://start.luma.fi/en/) in which students are making projects, was decided to be used as a tool in practice at the school level.

6.2 A practical stage in the co-design approach used

The co-design of the action plan for practice level in three phases of the Edelson's model (Edelson, 2002; Figure 1) was executed through a two-day design meeting, using mainly small group work and discussions. The main facilitators of the event were the project manager and the director of the program.

There were a lot of co-designers in the program: about 50 researchers, coordinators, project workers from the universities and a partner from industry facilitating by a national team (a director, a project manager, evaluation specialists, team leaders and chosen special researchers in science and technology education). They were divided into the teams of the chosen topics (see Section 6.1). The co-designers chose their own groups (e.g. math specialists participated in the math group). Each group had a group leader who facilitated discussion. As Roschelle & Penuel (2006) mentioned, co-design needs a bootstrapping event or process to catalyze the team's work and well-defined roles. After group discussions, the project manager summarised different ideas together and wrote a memorandum of the decisions (saved in the wiki platform) and shared it with all the co-designers via e-mail and also through the digital platform used.

Decision making was done in the co-design meeting, for example (i) about the digital platform (*Teams* selected) for co-design in detail, (ii) collecting evaluation materials from co-designers at school and (iii) a pre-questionnaire for co-designers at school before the first co-design meeting and timetables of the program. The co-designers suggested few suggestions for a digital platform (e.g. *Wiki*, *Teams* and *Facebook*). According to Andersen, Kanstrup & Yndigegn, (2018), there are many challenges with using technology for facilitating co-design. *Teams* were chosen because it is easy to use and teachers are using it a lot in Finland.

6.3 A "Bottom-up" stage in the co-design approach used

In the model, teachers are seen as active participants, professional key contributors and collaborators with researchers, as Roschelle & Penuel, (2006) found. The co-design of the previous action plan for supporting participating schools' and daycares' curricula based on empirical problem analysis (the needs of a school or a daycare), is seen as a critical phase of the co-design approach in the model. It is important to provide an opportunity to match the curriculum goals of teachers (cf. Tissenbaum et al, 2012; Kelly et al, 2019) and to increase reflections and ownership by a teacher (cf.

Roschelle, Penuel & Schechtman, 2006). According to Orland-Barak & Tillema (2006) also important for teachers, are: (i) shared an understanding of the challenge, (ii) a willingness to change one's own perspective, (iii) a commitment to participate in the dynamics of the group.

The main co-designers in this stage were teachers, the so-called LUMA mentors from each school or daycare. Their needs based on the curriculum of their school were taken into account in the co-design of the program. Altogether 160 voluntary schools and kindergartens (two members from each one) were chosen to the program by using the open call. A small group works between the co-designers were used in the meetings for facilitating the co-design approach.

The project workers from each LUMA Centre in a university (one for each chosen theme) and possible partners from industry were seen as facilitators for the co-design approach. The digital platform (*Teams*) for the co-design in detail was chosen by the teachers because it is easy to use and familiar to the teachers.

The co-design meetings were decided to be organised once a month during the development process. The schedule (only one year) is, however, quite tight. The way in which we fit the co-design approach to the school cycle is critical for success (cf. Roschelle & Penuel, 2006).

6.4 The co-design process as a tool for promoting teachers' professional development

The LUMA2020 program trusts the "bottom-up approach" for its success as in many earlier projects in which the DBR was found a useful way to promote teachers' or future teachers' professional development and growth (e.g. Vesterinen & Aksela, 2013; Aksela & Vihma, 2015; Aksela et al, 2016; Juuti, Lavonen & Meisalo, 2016). The co-design phases above describe the factors pointed out (Juuti et al, 2017): teacher-led, continuous (long-term), situated or connected to the classroom context, collaborative, and include reflective practices.

The systematic phases of the Edelson's design-based research model offer a learning environment, where teachers and all other participants can reflect and learn from one another, according to the 'learning community', especially in a cyclic development process of the model (Edelson, 2002; Figure 1). In practice, teacher educators as facilitators support teachers in the program to test the decided pilot model with their students, to collect research data and to reflect on the results in the monthly co-design meetings or through a digital platform that has been chosen together. Teachers as co-designers can also participate in writing the report and papers concerning the program or possible research facilitated by the teacher educators participating in the model.

7 Conclusions

The Edelson's design-based research model (DBR) used in the program can offer a systematic framework or a map for co-designing both the action plan and its implementation. Organizing the co-design approach within the model (Figure 1) through three main stages (see Sections 6.1, 6.2 and 6.3) with diverse multi-stakeholders (teachers, teacher educators, scientists or industry specialists as in the LUMA2020 program) could be fruitful for building relevant, novel practices in science education together if teachers are seen as active participants, key professional contributors and partners of researchers and teacher educators.

Guided face-to-face communication in the workshops or digital creative learning spaces as a medium for social, creative thinking could be useful for facilitating the co-design approach as in the LUMA2020 program. Familiar digital platforms can be used for planning, discussions, questions, sharing experiences and materials between co-designers within different stakeholders around Finland. Their role can be central for the co-design approach in practice when co-designers are far away from each other as in the LUMA2020 program.

The co-design model could help to bridge the newest research and innovations from industry into praxis for supporting the curriculum at school level and for promoting teachers' professional development by forming creative and diverse learning communities, in which all partners can learn from each other through sharing. It can also promote novel teacher training together with partners outside the universities (e.g. industry). Thus, the co-design approach implementation can offer a new kind of an educational model for both pre-service and in-service training. Teachers or possibly also future teachers can act as "researching teachers" in projects and learn through their reflection facilitated by the teacher educators.

The used Edelson's design-based research model (Figure 1) can be helpful for especially the novices in research and sponsors -to whom research and the process for research-based solutions are new things. The implementation of the co-design approach may increase relevant collaboration between schools, universities and the industry and commerce, promote collaboration between participants that are often unknown to each other (e.g. researchers, teacher educators, industry specialists, teachers, representatives from the educational administration and future teachers).

The limitation of this case study is the use of only a few documents as the main source. In order to better understand the successful co-design approach within design-based research with diverse multi-stakeholders, more research is needed to understand the different roles of the stakeholders (e.g. facilitators; teachers as co-designers), early design frames, design decision processes, creative learning spaces (e.g. digital platforms) for promoting the co-design approach and the views of its advantages and challenges for co-designers.

Acknowledgements

This article is based on my keynote presentation of the design-based research method and fruitful discussions on November 8th, 2018 in the symposium called "Contemporary approaches to research on mathematics, science, health and environmental education" at Deakin University, Australia. The best acknowledgements to the STEM research group at Deakin University, all LUMA collaborators (*https://www.luma.fi*) and especially my research team (*https://www.helsinki.fi/en/researchgroups/seco*) who has collaborated with me since the year 2005, and especially to a co-examiner, Topias Ikävalko.

References

Aksela, M. (2005). *Supporting meaningful chemistry learning and higher-order thinking through computer-assisted inquiry: a design research approach*. Helsinki: The University of Helsinki.

Aksela, M. K. (2010). Evidence-based teacher education: becoming a lifelong research-oriented chemistry teacher? *Chemistry education research and practice, 11*(2), 84–91.

Aksela, M., & Boström, M. (2012). Supporting Students' Interest through Inquiry-Based Learning in the Context of Fuel Cells. *Mevlana International Journal of Education, 2*(3), 53–61.

Aksela, M. K. & Vihma, L. J. (2015). Uudenlainen yhteisöllinen opettajankoulutus LUMA-ekosysteemissä elinikäisen oppimisen tukena (In English: Novel collaborative teacher education for long-life learning in the LUMA ecosystem). *LUMAT: Luonnontieteiden, matematiikan ja teknologian opetuksen tutkimus ja käytäntö 3*(6), pp. 711-720. https://www.lumat.fi/index.php/lumat-old/article/view/91

Aksela, M. K. (2015). Luma Centre Finland: The Finnish model to educate teachers in math, science and technology. In U. Heinichen, & O. J. Haupt (Eds.), *10 Jahre LeLa Jahrestagung* (pp. 16-18). Berlin: LernortLabor, Bundesverband der Schülerlabore.

Aksela, M., & Ikävalko V-M. (2016). How to promote relevant practical work in science education through a non-formal learning environment? *Science Education Research and Practical*

Work: A collection of invited papers inspired by the 23rd Symposium on Chemistry and Science Education held at the TU Dortmund University, May 26-28, 2016. Eds. Eilks, I, Markic, S. & Ralle, B. Aachen: Shaker, 2016.

Aksela, M. K., Vartiainen, J. L., Tuomisto, M., Turkka, J. S., Pernaa, J. I. S., & Tolppanen, S. (2016). Promoting Meaningful Science Teaching and Learning Through ICT in the Finnish LUMA Ecosystem. In H. Niemi, & J. Jia (Eds.), *New Ways to Teach and Learn in China and Finland: Crossing Boundaries with Technology* (pp. 255-278). Frankfurt am Main: Peter Lang.

Aksela, M., Oikkonen, J., & Halonen, J. (Eds.) (2018). *Collaborative Science Education at the University of Helsinki since 2003: New Solutions and Pedagogical Innovations for Teaching from Early Childhood to Universities*. Helsinki: Unigrafia Oy. https://www.helsinki.fi/sites/default/files/atoms/files/isbn-978-951-51-4089-0.pdf

Aksela, M. K., Pernaa, J. I. S., Blomgren, P. I. A., & Rautiainen, I. M. (2018). *Collaborative science education at the University of Helsinki: ChemistryLab Gadolin: a science lab as an inspiring environment for learning, development and research since 2008*. Helsinki: The University of Helsinki. https://www.helsinki.fi/sites/default/files/atoms/files/isbn-978-951-51-4506-2.pdf

Anderson, T., & Shattuck, J. (2012). Design-based research a decade of progress in education research? *Educational researcher, 41*(1), 16–25.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *Journal of the Learning Sciences, 13*(1), 1–14.

Baumfield, V. & Butterworth, M. (2007). Creating and translating knowledge about teaching and learning in collaborative school–university research partnerships: an analysis of what is exchanged across the partnerships, by whom and how. *Teachers and Teaching, 13*:4, 411–427.

Blomberg, J. L., & Henderson, A. (1990). Reflections on participatory design: Lessons from the Trillium Experience. Proceedings of CHI 90, (pp. 353-359). New York, NY:ACM.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences, 2*(2), 141–178.

Brown, A., & Campione, J. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles, and systems. In L. Schauble & R. Glaser (Eds.), *Innovations in learning: New environments for education* (pp. 289-325). Mahwah, NJ: Lawrence Erlbaum Associates.

Clements, D. H. & Battista, M. T. (2000). Designing Effective Software. In A. E. Kelly & R. A. Lesh. (Eds.) *Handbook of Research Design in Mathematics and science Education*, (pp. 761-776). New Jersey, Lawrence Erlbaum Associates.

Cobb, P. (2001). Supporting the improvement of learning and teaching in social and institutional context. In S. Carver & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 455-478). Cambridge, MA: Lawrence Erlbaum Associates.

Cognition and Technology Group at Vanderbilt. (1997). *The Jasper project: Lessons in curriculum, instruction, assessment, and professional development*. Mahwah, NJ: Erlbaum.

Collins, A. (1992). Towards a design science of education. In E. Scanlon & T. O'Shea (Eds.), *New directions in educational technology* (pp. 15-22). Berlin: Springer.

Collins, A. (1999). The changing infrastructure of education research. In E. Lagemann & L. Shulman (Eds.), *Issues in education research* (pp. 289-298). San Francisco: Jossey-Bass.

Dede, C. (2004). If design-based research is the answer, what is the question? A commentary on Collins, Joseph, and Bielaczyc; diSessa & Cobb; and Fishman, Marx, Blumenthal, Krajcik &

Soloway in the JLS special issue on design-based research. *Journal of the Learning Sciences, 13*(1), 105–114.

The Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, 5-8.

Edelson, D. C. (2002). Design research: What we learn when we engage in design. *The Journal of the Learning Sciences, 11*, 105–121.

Ehn, P. (1992). Scandinavian design: On participation and skill. In P. S. Adler & T. A. Winograd (Eds.), Usability: Turning technologies into tools (pp. 96-132). New York: Oxford University Press.

Finland and PISA, (2019). Finnish Ministry of Education and Culture. Finland and PISA. https://minedu.fi/en/pisa-en

Fischer, G., Giaccardi, E., Eden, H., Sugimoto, M., Ye,Y. (2005) Beyond binary choices: Integrating individual and social creativity. *International Journal of Human-Computer Studies, 63*(4-5), 482–512.

Hey, J.,Joyce, C. K. & Beckman, S. L. (2007). Framing Innovation: Negotiating Shared Frames During Early Design Phases. *Journal of Design Research, 6*(1), 77–99.

Huberman, A. M., & Miles, M. B. (1994). Data management and analysis methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 428-444). Thousand Oaks, CA, US: Sage Publications, Inc.

Kelly, A. E. (2003). Research as design. *Educational Researcher, 32*(1), 3–4.

Kelly, A. E., Lesh, R. A., & Baek, J. Y. (Eds.) (2008). *Handbook of design research in methods in education. Innovations in science, technology, engineering, and mathematics learning and teaching*. New York, NY: Routledge.

Kelly, N., Wright, N., Dawes, L., Kerr, J., & Robertson, A. (2019). Co-design for Curriculum Planning: A Model for Professional Development for High School Teachers. Australian *Journal of Teacher Education, 44*(7).

Krapp, A. & Prenzel, M. (2011) Research on Interest in Science: Theories, methods, and findings, *International Journal of Science Education, 33*(1), 27-50.

Juuti, K., & Lavonen, J. (2006). Design-based research in science education: One step towards methodology. *Nordina, 4*, 54–68.

Juuti, K. (2005). *Towards primary school physics teaching and learning: design research approach.* Helsinki: University of Helsinki.

Juuti, K., Lavonen, J. & Meisalo, V. (2016). Pragmatic Design-based Research –Designing as a Shared Activity of Teachers and Researchers. Iterative Design of Teaching-Learning Sequences: Introducing the Science of Materials in European Schools. Psillos, D. & Kariotoglou,P. (eds.). Dordrecht: Springer Science + Business Medisa, 35–46.

Lavonen, J. & Meisalo, V. (2002). Research-based design of learning materials for technology-oriented science education. *Themes in Education, 3*, 107–131.

Lavonen, J., Inkinen, J., Juuti, K., Salmela-Aro, K., Krajcik, J., & Schneider, B. (2017). The influence of an international professional development project for the design of engaging secondary science teaching in Finland. In M. K. Mhlolo, S. N. Matoti, & B. Fredericks (Eds.), *Book of Long Papers:25th Annual Meeting of the Southern African Association of Researchers in Mathematics Science & Technology Education (SAARMSTE)* (pp. 206-220). Johannesburg: SAARMSTE.

Lee, Y. (2008). Design participation tactics: The challenges and new roles for designers in the co-design process. *Co-design, 4*(1), 31–50.

Niemi, H., Lavonen, J., Kallioniemi, A., & Toom, A. (2018). The Role of Teachers in the Finnish Educational System: High Professional Autonomy and Responsibility. In H. Niemi, A.

Toom, A. Kallioniemi, & J. Lavonen (Eds.), *The Teacher's Role in the Changing Globalizing World: Resources and Challenges Related to the Professional Work of Teaching* (pp. 47-61). Leiden: Brill Sense.

Niemi, H. & Isopahkala-Bouret, U. (2015) Persistent Work for Equity and Lifelong Learning in the Finnish Educational System. *The New Educator, 11*:2, 130–145.

OECD (2015, 6) PISA 2015. PISA results in focus. http://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf

Osman Andersen, T., Kanstrup, A. M., & Yndigegn, S. L. (2018). Three living labs in Denmark: Challenges with co-design and implementation of health IT. I A. Bygholm, L. Pape-Haugaard, K. Niss, O. Hejlesen, & C. Zhou (red.), *Proceedings from The 16th Scandinavian Conference on Health Informatics 2018 Aalborg*, Denmark August 28–29, 2018 (s. 1-6). Linköping University Electronic Press. Linköping Electronic Conference Proceedings, Bind. 151

Noweski, C., Scheer, A., Büttner, N., von Thienen, J., Erdmann, J., & Meinel, C. (2012). Towards a paradigm shift in education practice: Developing twenty-first century skills with design thinking. In H. Plattner, C. Meinel, & L. Leifer (Eds.), *Design thinking research: Understanding innovation* (pp. 71-94). Berlin, Heidelberg: Springer-Verlag.

Penuel, W. R., Roschelle, J., & Shechtman, N. (2007). Teachers: An analysis of the co-desing process. *Learning, 2*(1), 51–74.

Pernaa, J. (Ed., 2013). *Kehittämistutkimus opetusalalla* (In English: Design-based research in education in English). Juva: Bookwell Oy.

Pernaa, J., & Aksela, M. (2013). Model-Based Design Research: A Practical Method for Educational Innovations. *Advances in Business-Related Scientific Research Journal*, *4*(1), 71–83.

Prahalad, C.K. & Ramaswamy, V. (2004). Co-Creation Experiences: The Next Practice in Value Creation. *Journal of Interactive Marketing*, *18*(3), 5-14.

Regan, E. & DeWitt, J., (2015). Attitudes, interest and factors influencing STEM enrolment behaviour: an overview of relevant literature. In E.K. Henriksen, J. Dillon and J. Ryder (Eds.), *Understanding student participation and choice in science and technology education.* (pp. 63-88). Dordrecht: Springer.

Research and development policy of the LUMA Centre Finland, (2018). LUMA Centre Finland, http://www.luma.fi

Richey, R. C., & Nelson, W. A. (1996). Developmental research. In D. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 1213–1245). London: Macmillan.

Rheinfrank, J. J., Hartman, W. R., & Wasserman, A. (1992). Design for usability: Crafting a strategy for the design of a new generation of Xerox copiers. In P. S. Adler & T. A. Winograd (Eds.), *Usability: Turning technologies into tools* (pp. 15-40). New York: Oxford University Press.

Roschelle, J., DiGiano, C., Koutlis, M., Repenning, A., Phillips, J., Jackiw, N., et al. (1999). Developing educational software components. *IEEE Computer, 32*(9), 50–58.

Roschelle, J., & Penuel, W. R., (2006). Co-design of innovations with teachers: Definition and dynamics. *Proceedings of the 7th International Conference on Learning Sciences*, 606–612.

Sanders, E. & Stappers, P.J. (2008) Co-creation and the new landscapes of design. *CoDesign, 4*(1), 5–18.

Sandoval, W. A., & Bell, P. (2004). Design-based research methods for studying learning in context: Introduction. *Educational Psychologist, 39*(4), 199–201.

Sandoval, W. (2014). Conjecture mapping: An approach to systematic educational design research. *Journal of the Learning Sciences, 23*(1), 18–36.

SECO, (2019). SECO (Science and Chemistry Education Collaboration) research group at University of Helsinki, Finland. https://www.helsinki.fi/seco

Shrader, G., Williams, K., Lachance-Whitcomb, J., Finn, L.-E., & Gomez, L. (2001). Participatory design of science curricula: The case for research for practice. Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA.

Sherin, M.G. (1998). Developing teachers' ability to identify student conceptions during instruction. In S. B. Berenson, K. R. Dawkins, M. Blanton, W. N. Coulombe, J. Kolb, K. Norwood, & L. Stiff (Eds.), *Proceedings of the Twentieth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 761-767). Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

Soloway, E., Guzdial, M., & Hay, K. E. (1994). Learner-centered design: The challenge for HCI in the 21st century. *Interactions, 1*(2), 36–47.

Schön, D.A. (1983). *The Reflective Practitioner*. New York, NY: Basic Books.

Taber, K.S. (2017). Science Education as a Field of Scholarship. In: Taber K.S., Akpan B. (Eds). *Science Education. New Directions in Mathematics and Science Education*. SensePublishers, Rotterdam.

Talbert, J.E. & M.W. McLaughlin (1999). Assessing the School Environment: Embedded Contexts and Bottom-Up Research Strategies," in S.L. Friedman & T.D. Wachs (Eds.) *Measuring Environment Across the Life Span,* (pp. 197-227). Washington DC: American Psychological Association.

Tissenbaum, M., Lui M., Slotta, J. & D. (2012). Co-Designing Collaborative Smart Classroom Curriculum for Secondary School Science. *Journal of Universal Computer Science, 18*(3), 327–352.

Zamenopoulos, T. and Alexiou, K. (2018). Co-Design as Collaborative Research in Facer, K and Dunleavy, K. (eds.) *Connected Communities Foundation Series*. Bristol: University of Bristol/AHRC Connected Communities Programme. https://connected-communities.org/wp-content/uploads/2018/07/Co-Design_SP.pdf

Wang, F. & Hannafin, M. J. (2004). Using design-based research in design and research of technology-enhanced learning environments. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.

van der Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (2006). Educational design research. London: Routledge. Whitehurst.

Vesterinen, V-M., & Aksela, M. (2013). Design of Chemistry Teacher Education Course on Nature of Science. *Science & Education*, *22*(9), 2193–2225.

Appendix 1

Table 3. The main principles given for the design process and the partners of the program (the text translated from the action plan written during the co-design phase 1 by the LUMA ecosystem).

The main principles for the design process	**The partners of the program or activity**
In the program, measures are carried out in two measure entities that in practice come together as one logical and intact program: A) **The sparking of motivation to learn and its support**, on the one hand in "classrooms" in different education levels and on the other hand during children, youth and families' free time. B) **The development of pre-service and in-service education for teachers** who teach math, environmental studies, biology, physics, chemistry and geography in different levels. **The program focuses on four focal points:**1. vocational and societal relevance in education 2. differentiated instruction, 3. the promotion of 'engineering skills' (technology education) and 4. creative, engaging work (e.g. embodied and drama-based learning) The common thread in the implementation of the LUMA 2020 program consists of a **differentiated viewpoint** that promotes mathematical and scientific literacy in a cross-currical way (uniting different fields and subjects, STEAM), in the spirit of multidisciplinary learning entities and theme learning according to the national core curricula. Both **'face-to-face' measures and virtual measures** are carried out in the program**,** which have effectiveness broadly all over Finland. **The program complements the national projects and actions that LUMA Centre Finland has already carried out** (see above). LUMA 2020 program complements the national **learning path for science and technology education**. The program promotes **educational continuity** of **educational and teaching staff** (early childhood educators, class teachers, special needs teachers, subject teachers, guidance counselors and vocational teachers). Spreading out occurs for example by educating future teachers and in-service teachers, in events organized by universities and other partners, in the form of academic and popular multimedia publication, and internationally also through researchers exchange and education export. Development can be published also for the scientific community and in the form of theses. The aim is to include this program as a part of partners' events such as events of educational organizations, already in Fall 2019 and especially in 2020. LUMA Centre Finland produces only such new data in its program that can be published using an open license. LUMA Centre Finland publishes all data from its program in a web platform that it operates and maintains this data at least until the program comes to its end, but depending on the possibilities data is available indefinitely.	**LUMA Centre Finland** (a network of 11 universities), where **13 regional LUMA Centres** operate. **University of Helsinki** is in charge of the administration of the network and it is **the responsible organizer of the LUMA 2020 program and other universities are its subcontractors.** 1)Nationally: In the implementation of the program, LUMA Centre Finland carries out development, education and marketing/communication collaboration with e.g. partners that are represented in the **national LUMA advisory board**. In addition, LUMA Centre Finland strengthens collaboration with **other national networks**. 2)Regionally/locally: **With Regional State Administrative Agencies, Education Division/ Education Consortiums of municipalities and individual learning environments** (units of early childhood education, units of primary and lower secondary education, units of upper secondary education, units of vocational education and training) and e.g. with **libraries, youth activities, sports clubs, hobby groups and parents' associations**. In individual learning environments, the aim is to get especially **tutors and guidance counselors** to participate actively in development. The aim is to also collaborate with university alumni already in the working life. LUMA Centre Finland strengthens especially the so-called **LUMA municipal network** that has been created in the LUMA FINLAND program during 2014-2016. In the concrete implementation of the measures of the program, the members of teams **collaborate** in the usual way **in their universities** with teacher educators, researchers and other staff as well as with basic degree students and postgraduate students.

The implementation of the program is evaluated throughout the entire process. Evaluation gives us information on how the implementation of the program is progressing concerning the aims set for development. Evaluation is divided into three areas: 1. Evaluating the extent/activity of operations and its quality 2. Evaluation of the effectiveness of operations 3. Evaluation of different actions (development of formal education, free-time activities and teacher's competences) In the first area of evaluation, the amounts of operational units (e.g. unit of early childhood education, primary and lower secondary school, upper secondary school, vocational institution), teachers, pupils, students and parents participating in the program are monitored during the entire process. **Measuring instruments are developed for evaluating the quality of the operations** and these instruments are based on surveys for guardians, the school and students, used in the PISA and TIMMS surveys, on the TALIS teacher survey and on national surveys such as the measuring instruments for entrepreneurship education. The approach makes it possible that the effectiveness of actions can be compared with the starting point (results of PISA, TALIS and TIMMS surveys). The evaluations rely on qualitative and quantitative sections. In addition, it's possible to carry out interviews for students and teachers. **Results** acquired from different areas of the evaluation program are **analyzed also vertically by combining results from various areas**. This is how a plausible bigger picture can be formed of the effectiveness of the different operational processes and its quality. This way it is possible to **reliably recognize good and unsuitable practices from one another and justifiable solutions can be made for developing actions and increasing the effectiveness**.	

Appendix 2

Table 4. The organization of the program and responsibilities of different partners and the stages of the program process the text translated from the action plan written during the co-design phase 1 by the LUMA ecosystem).

The organization of the program and responsibilities of different partners	**The stages of the program process**
The Finnish Ministry of Education and Culture assigns a **steering group** for the program that aligns the implementation of the program, and the following experts, the so-called core group are a part of the steering group. The **responsible leader** is the Director of LUMA Centre Finland from the University of Helsinki. **Experts** with their special expertise areas in the planning of the program, its implementation and evaluation include professors (science education, math education, technology education and the evaluation of education) and senior lecturers as representatives of vocational teacher education colleges. Experts participated for example by giving inserts/examples that guided learning. Those working alongside the program can consult experts and ask them about their views and they can receive support as the program is being implemented. In addition to experts, the program's **project manager** supports the responsible leader and directs the implementation of the program and is in charge of the national marketing/communication and collaboration with various partners. The mentioned experts and the project manager together form **the core group** of the program. The LUMA 2020 program is implemented nationally within the framework of the four themes of the StarT program. A team is formed for each theme, the so-called **theme team**. Each LUMA Centre participates in the implementation of each theme, in other words they participate in four theme teams; in the joint planning/forming of measures and in their individual geographical operations area, where they implement measures. In each of the LUMA Centres at least one worker works full time with the program or a couple of workers work part time on the program. The leaders of theme teams work in close collaboration together with the project manager. In addition, there is a team for the evaluation of the program, and it includes an evaluation specialist and the leaders of the four theme teams. The project manager is a part of all of the theme teams and the	**The iterative methodology of design-based research** is exploited in the collaborative planning, implementation and evaluation of the LUMA 2020 program. Measures selected for the program are driven forward with a developing way during the entire program, and novel, suitable solutions are produced to serve as everyday actions in different levels such as inspiring operations concepts and pedagogical approaches. At the same time the collaborative continuous culture of development in the LUMA ecosystem is strengthened and established. The systematic phases of design-based research offer a learning environment, **where all participants can reflect and learn from one another, according to the 'learning community'**: 1) mapping out needs together with participants (empirical problem analysis, the so-called needs-analysis), 2) mapping out new research information concerning the chosen topic from sciences, their learning and teaching (theoretical problem analysis), 3) setting the aims for development together with the participants based on steps 1 and 2, 4) planning a pilot model for the object of development (e.g. an activity, material) based on set aims, 5) testing the pilot model with the target group and refining the model based on received results (multiple steps), describing the development output and reporting and 7) spreading out new openings and solutions and offering education for these new topics. During the LUMA 2020 program, the development process is carried out at least in one cycle. The implementation of the program is divided into **two main phases**; development phase (from late Spring 2019 until Spring 2020) and spreading phase (from Spring 2020 until the end of 2020). In the first phase (from late Spring 2019 until Spring 2020) LUMA Centre Finland is going to build a national **network of LUMA development communities**. Learning communities operating in different levels (kindergartens, primary and secondary schools, upper secondary schools and vocational institutions) are encouraged to apply. The core group of the LUMA 2020 program chooses those learning communities that are accepted to the network of LUMA development communities, this is done based on the applications. Those learning communities selected to the network of development communities are engaged with their staff in collaborative development and at the same time in education, using **StarT as a tool**. In the academic year of 2019-2020, the StarT program is implementing learning projects in four various themes that those applying for the network of LUMA development communities can choose the most interesting themes for their own objects of development (and themes for projects): 1. Sustainable development (e.g. climate change, circular economy), 2) Mathematics around us (e.g. finances, art and statistics), 3. Technology around us (e.g. a moving device, artificial intelligence, robotics) and 4. My LUMA (topic free of choice, but one that is connected to LUMA subjects). The **theme teams plan** concrete measures in the framework of design-based research that can be made a part of current, existing national LUMA operations. Each theme team first familiarizes with already existing LUMA operating models and their material and models that are being planned currently. This ensures that the LUMA 2020 program **complements earlier LUMA programs and operations**. Together with **participating learning communities that have been selected** for the network of LUMA development communities, theme teams comprise aims for development and **together they test implementations of novel pedagogical solutions with learning communities**, especially on the level of early childhood education and upper secondary education.

evaluation teams. If needed, teams can consult the program's experts.	Each theme may consist of an educational package for all levels (also for primary and lower secondary education) offered to the learning communities and also **differentiated packages for early childhood education and upper secondary level**. The packages contain inserts/examples (in the form of videos) from the program's experts and other academic specialists. The inserts can be made e.g. from the viewpoint of the program's four focal points. Engaging education is carried out in practice with the model of blended learning, where (short) contact meetings and individual work through the online environment and especially interactions between colleagues working in different levels alternate. The staff of theme teams guide the development work for learning communities in **a low-threshold online environment** (e.g. in Facebook groups). In Fall 2019 a national welcome session is organized for the staff of LUMA development communities, which is organized simultaneously in LUMA Centres in different parts of Finland. After this session, LUMA Centres organize monthly 'face-to-face' meetings (2 hours each) until the late Spring of 2020. Alongside the regional StarT festivals in 2020, LUMA Centres offer educational program for the staff of LUMA development communities. The training for LUMA 2020 program's themes is in a central role in the national LUMA Days in June 2020. The staff of learning environments test pedagogical models in their teaching that are new for them, for example by instructing the students' project-learning and by carrying out their learning community's own StarT Day, possibly collect research data and reflect on the results in an online environment and/or in meetings together with the university personnel and with peers. They can also participate in writing research papers. LUMA development communities are encouraged to participate in regional StarT festivals in Spring 2020 with their students' project works. Models are formed for learning communities' StarT Days and regional StarT festivals for science and technology education activities for entire families and for working life collaboration. In addition, with these learning communities it is possible to form StarT science clubs for early childhood education and StarT clubs e.g. for the lower and upper secondary level. Here, also the staff of learning communities can test how to instruct project learning. **In the second phase (from Spring 2020 to the end of 2020), online-courses (at least 1-3) are created for themes.** With the help of these it is possible to promote operations of learning communities as well as the competences of educational and teaching staff more broadly and with long-term effect, because these can be exploited even after the LUMA 2020 program in teachers' pre-service education and in support of continuous development and as a part of support for the learning community from the StarT program. In addition, virtual club packages can be formed from the themes for small children with their families and for youth. The purpose is that the learning communities belonging in the network of LUMA development communities operate in Fall 2020 and even after that in the roles of LUMA peer mentors for other learning communities nearby. The contents of the LUMA 2020 program are in a central role in 2020 in all operations that LUMA Centre Finland monitors in the forums of teacher education in LUMA subjects **aimed for university personnel**, research on teaching and development. Through the program, the aim is to include more aspects from universities and also teacher educators from early childhood education, teacher training schools and from universities of applied sciences. Through the broadening of the forum, tripartite and four-party collaboration between operators of teacher education are promoted.

A systematic review of educational design research in Finnish doctoral dissertations on mathematics, science, and technology education

Daranee Lehtonen, Anne Jyrkiäinen and Jorma Joutsenlahti

Faculty of Education and Culture, Tampere University, Finland

Since educational design research (EDR) was introduced to educational research at the beginning of the 1990s, it has gained recognition as a promising research approach that bridges the gap between research and practice in education. This paper aims to investigate how EDR has been utilised and developed and which challenges it has faced by systematically reviewing 21 Finnish EDR doctoral dissertations on mathematics, science, and technology education published between January 2000 and October 2018. The findings indicate that all dissertations yielded practical and theoretical contributions. Moreover, common EDR characteristics, including the use of educational problems in practice as a point of departure, research in real-world settings, evolution through an iterative process, development of practical interventions, and refinement of theoretical knowledge, were found in all dissertations. Most of the doctoral researchers were confronted with challenges, such as high demand for EDR with limited resources and difficulties associated with multidisciplinary teamwork. However, the dissertations were diverse in terms of research contexts, practical educational problems, research outcomes, research methodologies, scale, and collaboration. This systematic review not only enhances the understanding of the utilisation, development, and challenges of EDR but also provides implications for future EDR.

Keywords
educational design research, doctoral dissertation, mathematics education, science education, technology education

Correspondence
daranee.lehtonen@tuni.fi

DOI
https://doi.org/10.31129/LUMAT.7.3.399

1 Introduction

Since educational design research (EDR) was introduced to educational research at the beginning of the 1990s, it has gained recognition as a promising research approach that bridges the gap between theoretical research and practice in education. Globally, EDR is still developing (Easterday, Lewis, & Gerber, 2017), as it is relatively young compared to other research approaches in education (Bell, 2004; Ørngreen, 2015). Over the past three decades, researchers have conducted EDR from a variety of theoretical perspectives and traditions for various purposes and contexts using different research methods (Bell, 2004; Prediger, Gravemeijer, & Confrey, 2015). While they have provided evidence supporting the usefulness of EDR, some have critiqued its limitations and challenges.

To better understand how EDR has been utilised and developed and which challenges it has faced, we systematically reviewed EDR studies conducted in the context of mathematics, science, and technology education at all levels. This lens was chosen for two reasons. First, it is likely that EDR is conducted differently in different educational fields, and therefore examining its application in specific fields may help refine the understanding of how to carry out EDR (McKenney & Reeves, 2012). Second, EDR has been adopted in a growing body of research on mathematics, science, and technology in education (Anderson & Shattuck, 2012; Prediger et al., 2015; Zheng, 2015).

2 Educational design research (EDR)

2.1 Overview of EDR

In this paper, we use the term *educational design research* to describe a research approach that is also known as *design experiments*, *design research*, *design-based research*, and *development (al) research*. EDR uses educational problems in practice as a point of departure and seeks to develop practical solutions to improve educational practices and advance usable knowledge through iterative processes in real-world settings (McKenney & Reeves, 2019; Plomp, 2013).

The manifold studies on EDR differ in terms of goals, forms, processes, outcomes, and other aspects (e.g., Bell, 2004; Plomp, 2013; Prediger et al., 2015). In addition, scholars have defined EDR in a variety of ways. Table 1 provides examples of EDR characteristics proposed by Anderson and Shattuck (2012); Cobb, Confrey, diSessa, Lehrer, and Schauble (2003); Juuti and Lavonen (2006); McKenney and Reeves (2019); and Wang and Hannafin (2005). Nevertheless, there are some commonalities among the definitions: intervention in real-world settings to improve practices, evolution through iterative cycles, development of practical solutions (i.e., interventions), and refinement of theoretical knowledge.

Table 1. Variants of educational design research (EDR) characteristics proposed by different scholars.

Title	Characteristics of EDR	Reference
Design-based research	1. Situated in real educational contexts 2. Focusing on the design and testing of interventions 3. Utilising mixed methods 4. Involving multiple iterations 5. Entailing partnership between researchers and practitioners 6. Providing design principles 7. Different from action research 8. Having a practical impact on practice	Anderson and Shattuck, 2012, pp. 16–18
Crosscutting features of design experiments	1. Developing theories about the learning process and ways to facilitate that learning 2. Interventionist: bringing about educational innovation 3. Prospective and reflective 4. Iterative cycles of intervention and revision 5. Practice orientated	Cobb, Confrey, diSessa, Lehrer, and Schauble, 2003, pp. 9–11
Features of the design-based research	1. Iterative process 2. Developing usable artefacts 3. Rendering novel educational knowledge	Juuti and Lavonen, 2006, pp. 59–63
Features of the design research process	1. Theoretically oriented 2. Interventionist: developing solutions informed by existing knowledge, testing, and participants 3. Collaborative: working in collaboration with others 4. Responsively grounded process 5. An Iterative process of investigation, development, testing, and refinement	McKenney and Reeves, 2019, pp. 12–16
Characteristics of design-based research	1. Pragmatic: refining theory and practice 2. Grounded in relevant research, theory, and practice 3. Interactive: working together with participants; an iterative cycle of analysis, design, implementation, and redesign; and flexible when necessary 4. Integrative: using mixed research methods 5. Contextual research results and generated design principles	Wang and Hannafin, 2005, p. 8

Descriptions of phases of EDR differ between scholars (cf. Cobb et al., 2003; Easterday et al., 2017). According to Plomp (2013), there are three main phases: (1) preliminary research (i.e., literature research, needs and context analysis, and theoretical framework development), (2) the development phase (i.e., the iterative design phase), and (3) the assessment phase (i.e., the summative evaluation of the intervention and recommendations for improvement; cf. McKenney & Reeves, 2019, who described the initial phase, design phase, and evaluation). McKenney and Reeves (2019) divided EDR into cycles of different sizes: single subcycle, multiple subcycles, and overall design research project. A single subcycle is the completion of one of the three main phases (i.e., preliminary research, development, or assessment). Multiple subcycles consist of several subcycles, but not as many as the whole EDR project. An overall design research project can range from one multiple subcycle that consist of three subcycles of each phase to several multiple subcycles.

EDR contributes to both practice and theory. In terms of its practical contribution, EDR uses an iterative process of design, assessment, and redesign in authentic contexts to develop an intervention to solve an educational problem (McKenney & Reeves, 2019). Additionally, according to Edelson (2002), EDR can help to develop three types of theory: *domain theories*, *design frameworks*, and *design methodologies* (cf. Plomp, 2013). Domain theories describe real-world phenomena and the outcomes of design implementation; design frameworks describe the characteristics of successful solutions to the problem in the studied context; design methodologies provide guidelines for successfully achieving the research aims.

2.2 EDR challenges and recommendations

Scholars have addressed several challenges of EDR and provided recommendations for how to overcome them. First, the triangulation of data sources, data collection methods, data types, theories, and evaluators is recommended to better understand complex real-world phenomena and enhance the reliability and validity of EDR (e.g., Design-Based Research Collective [DBRC], 2003; McKenney & Reeves, 2019). Nevertheless, triangulation and the iterative nature of EDR usually lead to *over methodologisation*—that is, the collection and analysis of excessive amounts of data—which sometimes many not lead to adequate results (e.g., Brown, 1992; Dede, 2004). Second, EDR researchers often take on multiple roles (e.g., researcher, designer, implementor, and evaluator of the intervention), which may lead to conflicts of interest (e.g., Plomp, 2013). Triangulation of researchers can enhance the objectivity of EDR (Plomp, 2013). Third, several EDR studies tend to be *under conceptualised*, as they lack a profound theoretical foundation and do not seek to provide theoretical contributions (e.g., Dede, 2004). Therefore, EDR should not only provide solutions to problems but also yield a variety of theories, particularly theories related to the design process (McKenney & Reeves, 2019). Fourth, a multidisciplinary collaboration among various experts from relevant fields is recommended for ensuring the feasible and successful development of solutions to complex educational problems (e.g., Wang & Hannafin, 2005). However, multidisciplinary teamwork requires, for example, a shared understanding among team members, strong group cohesion, and respect for others, and thus teamwork can be tiresome and contentious (McKenney & Reeves 2019). Fifth, the involvement of various participant groups that are relevant to the implementation of the intervention (e.g., teachers, students, and organisations) is advised to better understand complex authentic contexts and enhance respondent triangulation (McKenney & Reeves, 2019; Ørngreen, 2015). Sixth, rather than refining only one design idea, working with alternative designs and exploring solutions is recommended to ensure that the proposed intervention is the best solution to the problem (McKenney & Reeves, 2019; Ørngreen, 2015). Finally, Kelly (2013) proposed that, as EDR requires the investment of considerable resources, EDR should be employed only when truly needed, such as when facing a challenging educational problem with no satisfactory solution.

2.3 Previous reviews of EDR

Previous studies have investigated the utilisation and progress of EDR and other relevant issues with various focuses and review processes.

Anderson and Shattuck (2012) reviewed and defined the characteristics of EDR, including interventions in real educational contexts, a focus on the design and testing of a significant intervention, the use of mixed methods, multiple iterations, a collaborative partnership between the researcher and practitioners, the provision of design principles, differences from action research, and practical impact. The authors also conducted a review of the 47 most cited EDR articles from 2002 to 2011. Quantitative and qualitative content analyses were conducted to investigate geographic, disciplinary, and curricular focuses and the interventions, iterations, and outcomes of the articles. They found that design research was increasingly employed in educational contexts and that the majority of studies were conducted in North America. The most commonly studied subject was science; the main context was K–

12; and most interventions involved technology. Thirty-one articles were empirical studies that were part of a multi-iterative research project. All of the empirical studies involved were either technological and instructional design interventions or instructional methods, models, and strategies. Typically, mixed methods were employed. Most focused on furthering theoretical knowledge and developing applications to improve learners' learning outcomes or attitudes. Although the results of their review affirmed the great promise of EDR due to its integration of educational theory and practice, Anderson and Shattuck (2012) argued that work still needs to be done regarding educational innovations. Moreover, they recommended that future reviews perform a more detailed investigation of the full text of articles and investigate a broader set of articles. Their characterisation of EDR has been cited numerous times.

According to McKenney and Reeves (2013), most of the EDR characteristics defined by Anderson and Shattuck (2012) are similar to those reported by other authors. However, McKenney and Reeves identified that departure from a problem is an important characteristic of EDR that is missing from Anderson and Shattuck's (2012) list. Moreover, they criticized Anderson and Shattuck's (2012) systematic review for its limited search terms (*design-based research* and *education*), narrow dataset (i.e., only the most cited articles), and the use of only abstracts for a number of analyses (McKenney & Reeves, 2013). They called for the use of diverse search terms, an adequate dataset, and in-depth analyses of full texts to assess EDR progress in future studies (McKenney & Reeves, 2013).

Kennedy-Clark (2013) provided an overview of EDR as well as emphasised Plomp's (2007) three phases of EDR (i.e., initial, prototyping, and assessment phases) and the contribution of iterative cycles to the development of design principles and the refinement of theories. Furthermore, she investigated how EDR characteristics were used in doctoral dissertations by critically reviewing six education dissertations utilising EDR that were published by different institutions in Australia, Europe, Africa, and North America from January 2000 to January 2013. Her search terms included *design research*, *design-based research*, *education*, *phases*, *cycles*, and *iteration*. The research contexts (i.e., teaching subjects and education levels), focuses, and duration of data collection cycles varied among the dissertations, but they all utilised mixed methods for data collection. Conducting iterative data collection phases, engaging with several expert groups, testing designs with different participation groups, and being flexible and adaptive appeared to assist the researchers in reflecting on their research, understanding the educational problem, and avoiding overstated claims and conclusions. Finally, Kennedy-Clark's review demonstrated that the use of iterative design and development cycles or micro phases could increase the reliability and trustworthiness of research.

As researchers interested in EDR, we appreciate Kennedy-Clark's in-depth review of the potential benefits of EDR for education dissertations. However, the method was not sufficiently elaborated, and no overview of the information in the dissertations was provided. Revisiting the original article (Kennedy-Clark, 2013), Kennedy-Clark (2015) highlighted that researchers tend to concentrate on publishing their research findings and neglect to report their research methodologies. Therefore, there is a need for further investigation of how researchers employ EDR in their studies (Kennedy-Clark, 2015).

Zheng (2015) noted that applications of EDR do not appear to live up to expectations. She investigated empirical studies that adopted EDR through a systematic review of 162 journal articles published between 2004 and 2013 and quantitative content analysis of the selected EDR studies in terms of demographics, research methods, intervention characteristics, and research outcomes. The findings show that higher education was the most common sample group, and natural science was the most commonly studied learning domain. Qualitative methods were most often adopted, mixed methods were the second most popular, and solely quantitative methods were not used in any studies. Nearly all studies collected miscellaneous data, including interviews, questionnaires, and notes; and most performed technological interventions. More than half of the studies designed, developed, and redesigned educational interventions in only one iteration cycle. Although the majority revised their interventions, only approximately half of the studies reported how they did so. Moreover, most studies relied heavily on measurements of learners' cognitive outcomes. Based on her findings, Zheng (2015) proposed that there is a need for EDR studies to apply multiple iterations and new approaches that pay more attention to the design process.

We value her work for its thorough review of a large number of EDR studies and because it improves the understanding of the EDR landscape over the past decade. Nevertheless, a more detailed qualitative analysis would have complimented her quantitative analysis and contributed to an even deeper understanding of the selected studies. Zheng (2015) recognised the shortcomings of her research and recommended more deliberate investigation and analysis of design activities and their functions.

3 Methodology

3.1 Dissertation search and selection

To investigate how EDR has been employed in research on mathematics, science, and technology education and which challenges have confronted EDR researchers, we conducted a systematic review based on the recommendations of Anderson and Shattuck (2012), Kennedy-Clark (2015), McKenney and Reeves (2013), and Zheng (2015; see Section 2.3). Our data was collected from Finnish doctoral dissertations on mathematics, science, and technology education published between January 2000 and October 2018. We chose dissertations as our dataset because they report all iterative phases of the completed research, unlike articles, which often report only specific phases of research. We focused on Finnish dissertations because, as researchers in Finland, we expected our familiarity with the Finnish education system and practices to assist our review. It was not feasible to review all related dissertations completed at all Finnish universities because each university's repository uses a different database system, and there is no shared database containing all Finnish dissertations. Therefore, we decided to retrieve our data from the institutional repositories of the five Finnish universities that awarded the most qualifications and degrees in 2014: the University of Helsinki, University of Jyväskylä, University of Oulu, University of Tampere, and University of Turku (Official Statistics of Finland, 2015). The repository of the University of Eastern Finland, which provided the fourth most qualifications and degrees in 2014 and where a number of EDR dissertations have been completed, did not support the use of search terms for data retrieval. We also tried to retrieve dissertations of the University of Eastern Finland from Finna, a collection of search services providing access to material from Finnish university libraries. However, the Finna portal did not support a full-text search, which we used in our systematic review. Thus, we excluded the University of Eastern Finland and included the University of Tampere instead. Although our list of dissertations is not comprehensive, we believe that it provides an overview of the various dissertations published in Finland.

Our search terms included different terminologies that have been used to describe EDR in both English (*design research*, *design-based research/design based research*, *development research/developmental research,* and *design experiments*) and Finnish (*design-tutkimu*/suunnittelututkimu**, *design-perustai*/design-perustei*/suunnitteluperustai*/suunnitteluperustei**, *kehittämistutkimu*,* and *design-eksperiment**). The initial search resulted in 625 dissertations. One of the authors and a research assistant screened these results using the following inclusion criteria: (1) at least one of the search terms is visible in the English or Finnish title, abstract, or keywords and (2) the full text is openly available digitally. After applying these criteria, 55 dissertations remained. Each of the authors independently read one-third of this list according to our own interests and expertise. Thereafter, we jointly decided to exclude dissertations that did not utilise EDR as a strategy of inquiry, leaving 49 dissertations. At the beginning of this research, we decided not to use search terms similar to *mathematics* OR *science* OR *technology* AND *teach** OR *learn** OR *class** to locate all dissertations on mathematics, science, or technology education because doing so would not be possible. Instead, we carefully read the remaining EDR dissertations, identified which dissertations concerned mathematics, science, and technology education, and jointly excluded dissertations in fields other than mathematics, science, and technology education, such as other taught subjects (e.g., language, design, and nursing), skill and competence development, teaching and learning support, and learning environments in general.

3.2 Dataset

After the final screening process, the full texts of 21 EDR dissertations (10 in English and 11 in Finnish; 18 monographs and 3 article-based dissertations) on mathematics, science, and technology education from three universities (the University of Helsinki, University of Jyväskylä, and University of Oulu; n = 14, 6, and 1, respectively) remained for statistical and content analysis. Table 2 presents the number of EDR dissertations on mathematics, science, and technology education and on other educational domains by the university during the periods of 2000–2009 and 2010–2018.

Table 2. Frequency of EDR dissertations by the university and educational domain.

Educational Domains	University of Helsinki		University of Jyväskylä		University of Oulu		University of Tampere		University of Turku	
Year	2000 – 2009	2010 – 2018	2000 – 2009	2010 – 2018	2000 – 2009	2010 – 2018	2000 – 2009	2010 – 2018	2000 – 2009	2010 – 2018
Mathematics, science, and technology	4	10	3	3	0	1	0	0	0	0
Other	2	8	2	5	2	3	2	3	0	1
Total	6	18	5	8	2	4	2	3	0	1

Among the EDR dissertations on mathematics, science, and technology education, those of Aksela (2005) and Juuti (2005) were the first two published at the University of Helsinki, that of Leppäaho (2007) was the first at the University of Jyväskylä, and that of Oikarinen (2016) was the only one published at the University of Oulu. Altogether, there were 19 supervisors for the 21 dissertations. Aksela, who completed her EDR dissertation in 2005, supervised nine dissertations (43%), while Lavonen supervised six dissertations (29%).

3.3 Data analysis

After the final screening, each author coded one-third of the dissertations using a jointly constructed coding table. The coding categories were initially based on the previous literature, but we regularly discussed and modified existing categories and added relevant categories during the coding to best answer our research questions.

We coded the dissertations according to the following categories: (1) use of EDR terms and theoretical frameworks, (2) research contexts (i.e., educational sectors, settings, and domains), (3) educational problems in practice and research outcomes, (4) research methodology (i.e., research methods, data collection methods, and data sources), (5) scale, collaboration, and researcher's roles, (6) EDR process (i.e., phases of EDR, iterations, alternative design interventions, and issues during development of the intervention), and (7) EDR challenges. After the coding, we analysed the coded data quantitatively and qualitatively. Our findings are presented according to these seven categories in tables, figures, and descriptive analyses in the following section.

During the study, we strived to enhance the validity and reliability of our study by performing a precise research process, making joint decisions, crosschecking our data and analysis, consulting the literature for interpretations of the data, and comparing our research results to previous studies.

4 Results

4.1 Use of EDR terms and theoretical frameworks

EDR is referred to by a variety of names, and different scholars define it as having different goals, characteristics, and processes. Thus, investigating how EDR terms and theoretical frameworks have been used in dissertations published during the last two decades improves the understanding of how EDR is utilised and developed.

Of the four terms in each language used for our dissertation search, only three — design research, design-based/design based research, and development/developmental research in English and design-tutkimu*/suunnittelututkimu*, design-perustai*/design-perustei*/suunnitteluperustai*/suunnitteluperustei*, and kehittämistutkimu* in Finnish — appeared in the titles, abstracts, or keywords of the 21 dissertations. The dissertations did not apply a uniform format: while all of the dissertations included English versions of the title and abstract, only 18 included Finnish versions. We counted the appearance of each term only once per dissertation. Vartiainen (2016) used two terms in her English abstract, and Hassinen (2006) used two terms in her Finnish abstract. Thus, we also included them in our data (English: n = 22; Finnish: n = 19).

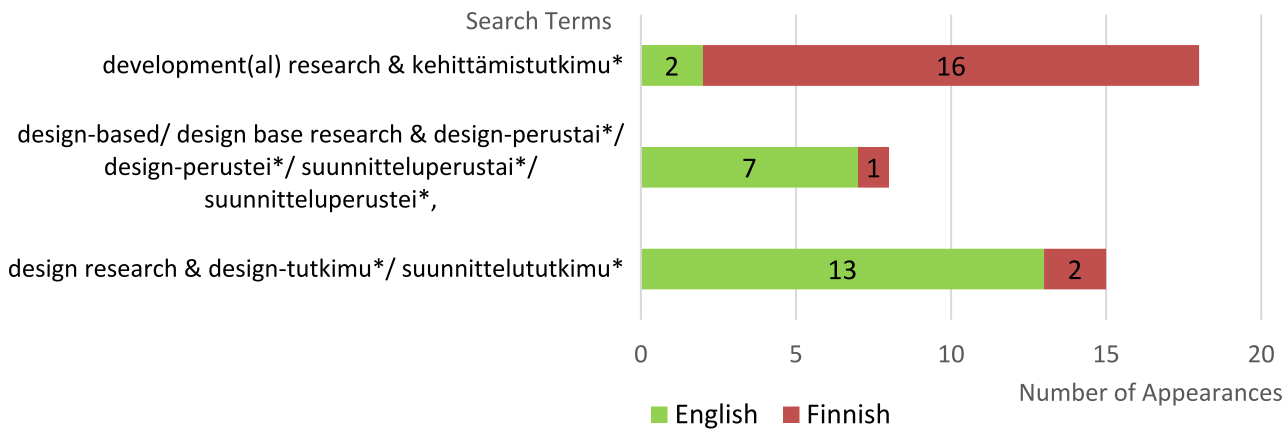


Figure 1. Frequency of search terms appearing in dissertation title, abstract, or keywords (English: n = 22; Finnish: n = 19).
Note: Only 18 dissertations had Finnish abstracts. One researcher used two terms in the English abstract, and another used two terms in the Finnish abstract.

Figure 1 shows the frequency of each search term in the English or Finnish titles, abstracts, or keywords of the dissertations. The most commonly used English term was ‘design research’, which appeared in 13 dissertations (69%), followed by ‘design-based/design based research’ in 7 dissertations (32%). The most commonly used Finnish term was ‘kehittämistutkimu*’ (development/developmental research), which appeared in 16 dissertations (84%). Interestingly, for 13 of the 18 dissertations (72%) that provided the title and abstract in both languages, the English and Finnish terms were not consistent. These dissertations used ‘kehittämistutkimu*’ (development/developmental research) in their Finnish titles or abstracts but either ‘design research’ or ‘design-based/design based research’ in their English titles or abstracts.

Our search terms appeared in the titles of 12 dissertations (57% of the 21 dissertations). Of these, six (50%) included the search terms in their primary titles, such as “**Design-Based Research** of a Meaningful Nonformal Chemistry Learning Environment in Cooperation with Specialists in the Industry” (Ikävalko, 2017) and “A **Design Research**: Problem and Inquiry Based Higher Education of Chemistry” (Rautiainen, 2012).

The comprehensiveness with which EDR theoretical frameworks were presented in the methodology sections of the dissertations varied from relatively superficial to exceedingly thorough. To investigate the use of these theoretical frameworks, we focused on the main EDR literature cited in the dissertations’ methodology sections, such as those regarding the principles, key characteristics, and processes of EDR. We found that early EDR works (e.g., Brown, 1992; Edelson, 2002; DBRC, 2003) and recent works (e.g., Anderson & Shattuck, 2012; McKenney & Reeves, 2019) were used as the main theoretical frameworks. The most cited article was that of Edelson (2002), which described the three types of theories (i.e., domain theories, design frameworks, and design methodologies) that can guide EDR. This article was cited in 18 dissertations (86%). The next most cited article was that of the DBRC (2003), which identified five characteristics of good design-based research and provided recommendations on how to increase the reliability and validity of EDR. This article was cited in 10 dissertations (48%). Of the Finnish EDR literature, Juuti and Lavonen’s (2006) article concerning the three pragmatic features of EDR was cited by nine dissertations (43%).

4.2 Research contexts

To obtain an overview of the authentic educational contexts in which the EDR dissertations were conducted, we examined their research contexts, including the educational sector (i.e., educational levels based on the Finnish educational system), setting (i.e., formal education vs. nonformal education), and domain (i.e., teaching and learning subjects).

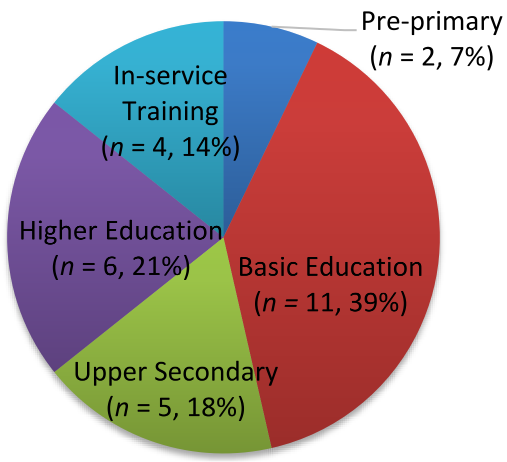


Figure 2. Frequency of educational sectors examined by the dissertations (n = 28)

Note: Five dissertations were carried out in more than one educational sector.

All of the dissertations were conducted in real-world educational contexts, and five were carried out in more than one educational sector. We included all of these sectors in our data (n = 28). Figure 2 shows a pie chart of the various educational sectors examined by the dissertations. Basic education (Grades 1–9; n = 11, 39%) was the most studied educational sector in the dissertations, while pre-primary school (n = 2; 7%) was the least.

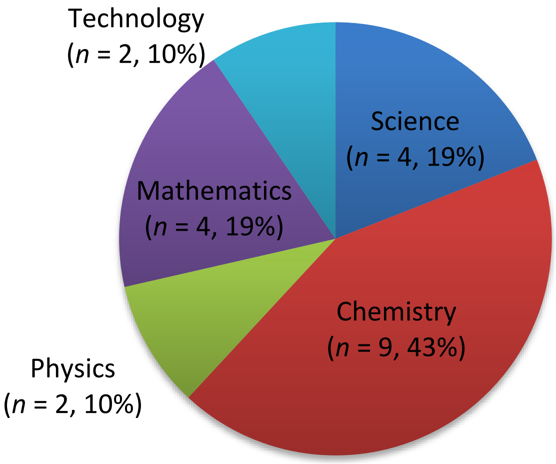


Figure 3. Frequency of the educational domains on which the dissertations focused (n = 21)

Note: The education domains were categorised based on the vocabulary used in the dissertations.

The majority of the 21 dissertations (n = 14, 67%) were conducted in a formal educational setting leading to formal qualifications, while the others were conducted in either a nonformal setting (n = 3, 14%) or in both types of settings (n = 4, 19%). The research interventions were conducted in various educational domains. Some researchers described these domains in a general way (e.g., science, mathematics, or technology in education), while others referred to specific subjects (e.g., chemistry and physics). We categorised our data accordingly. Moreover, we included upper secondary school statistics for mathematics, which is in line with the Finnish national core curriculum. Figure 3 illustrates that the most common domain was chemistry (n = 9, 43%), followed by science in general (n = 4, 19%) and mathematics (n = 4, 19%).

Table 3. Three dissertations serving as examples of variations in the research contexts of the dissertations

Research contexts	**Rukajärvi-Saarela (2015)**	**Ekonoja (2014)**	**Vartiainen (2016)**
Educational sector	Pre-service teacher education and in-service teacher training	Lower and upper secondary education	Pre-primary education (ages 3–6)
Educational setting	Formal and nonformal education	Formal education	Nonformal education
Educational domain	Primary school chemistry teaching	Teaching information and communication technology	Science club for small children

In sum, the dissertations were conducted in various research contexts (i.e., educational sectors, settings, and domains). Table 3 illustrates the differences in the research contexts using three dissertations as examples.

4.3 Educational problems in practice and research outcomes

All dissertations took at least one of four types of practical educational problems as a point of departure. Two dissertations took two types of problem as a point of departure; thus, we also included them in our data (n = 23). Figure 4 shows that the most common problem (n = 11, 48%) was students' lack of motivation and interest (e.g., Vartiainen, 2016), low performance (Hassinen, 2006), or deficient understanding (e.g., Oikarinen, 2016). The second most common problem (n = 7, 30%) was a lack of teaching and learning materials (e.g., Hongisto, 2012) or challenges in adapting to a new teaching and learning environment (e.g., Nieminen, 2008). The third type of problem (n = 3, 13%) was a teachers' deficient understanding and pedagogical skills (e.g., Juntunen, 2015). The last type (n = 2, 9%) concerned changes in a new curriculum (e.g., Kallunki, 2009).

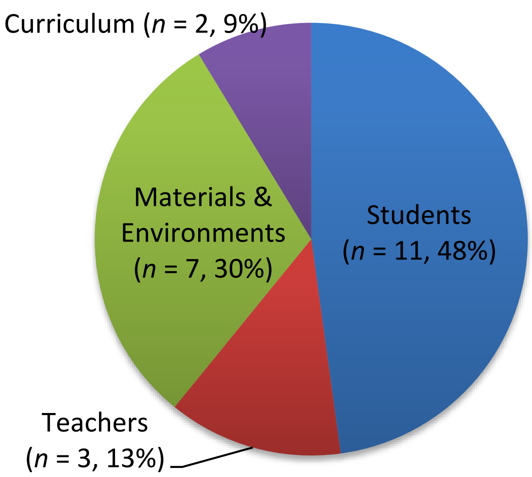


Figure 4. Types of practical educational problems that the dissertations took as points of departure (n = 23)

Note: Two dissertations took two problems as points of departure.

With regard to the practical contributions of the dissertations, various educational interventions were developed to respond to educational challenges in practice. Kallunki (2009) developed a teaching model and a learning environment, and we included both in our data (n = 22). The most common type of intervention involved teaching and learning environments (n = 10, 45%), such as a virtual science club

(Vartiainen, 2006) or a chemistry information and communication technology (ICT)-based learning environment (Pernaa, 2011). Another major type concerned teaching and learning concepts or models (*n* = 9, 41%), such as new chemistry teaching concepts for sustainability education (Juntunen, 2015) or a teaching model for algebra (Hassinen, 2006). Teaching and learning materials (*n* = 3, 14%), such as textbooks and electronic learning materials for teaching ICT (Ekonoja, 2014), were also developed.

We also investigated the theoretical contributions of the dissertations. Figure 5 shows that the majority of the dissertations (*n* = 15, 71%) developed all three types of theory (i.e., domain theories, design frameworks, and design methodologies) described by Edelson (2002). Nonetheless, only 11 of 15 developed all these theories thoroughly (e.g., Vartiainen, 2016). The remainder (*n* = 6, 29%) only developed domain theories and design frameworks (e.g., Tomperi, 2015) or domain theories and design methodologies (Leppäaho, 2007).

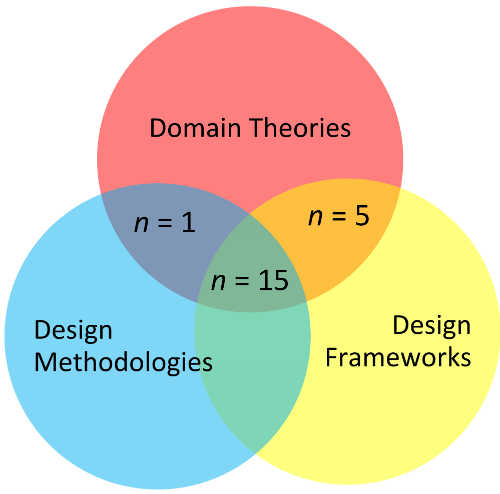


Figure 5. Venn diagram illustrating the theoretical contributions of the dissertations (*n* = 21)

4.4 Research methodology

The way in which EDR projects are conducted plays an important role in the success and reliability of those projects. Research triangulation is highly recommended to ensure the quality of EDR. Therefore, we examined how the triangulation of research methods, data collection methods, and data sources was implemented in the dissertations.

We coded the research methods as qualitative, quantitative, and mixed methods (see e.g., Creswell & Creswell, 2018). Fourteen dissertations (67%) gathered and analysed data with mixed methods (i.e., both qualitative and quantitative methods), while the remainder ($n = 7$, 33%) used only qualitative methods. None were conducted with only quantitative methods. Nevertheless, some of those dissertations that adopted mixed methods did not utilise qualitative and quantitative methods equally. For example, Ratinen's (2016) dissertation consisted of three substudies, only the first of which adopted mixed methods (i.e., a qualitative and quantitative questionnaire).

The dissertations used various methods to collect empirical data. The most common data collection methods were observation and questionnaires (each of which was used by 15 dissertations), followed by written documents, such as essays, diaries, and reports (which were used by 14 dissertations), and then interviews and group interviews (used by 13 dissertations). Some dissertations used tests and exams (e.g., Nieminen, 2008), tasks and exercises (e.g., Juntunen, 2015), and design intervention analysis (e.g., Pernaa, 2011). With regard to data sources, approximately half (11 of 21) of the dissertations collected data from both students and teachers, while the other half ($n = 10$) collected data from only students or only teachers. Additionally, several dissertations collected data from sources other than students and teachers; for example, Ikävalko (2017) collected data from company specialists, and Vartiainen (2016) collected data from parents.

In addition to investigating the dissertations' data collection methods and sources, we investigated how they collected data with multiple methods and from multiple sources to enhance their research triangulation. The number of data collection methods used in each dissertation ranged from one (Hongisto, 2012) to seven (Juuti, 2005), and the majority used three ($n = 7$, 33%) or four ($n = 5$, 24%). The number of data sources used in each dissertation varied from one (e.g., Rukajärvi-Saarela, 2015) to five (Tuomisto, 2018). Most of the researchers collected their data from one ($n = 7$, 33%) or two sources ($n = 10$, 48%).

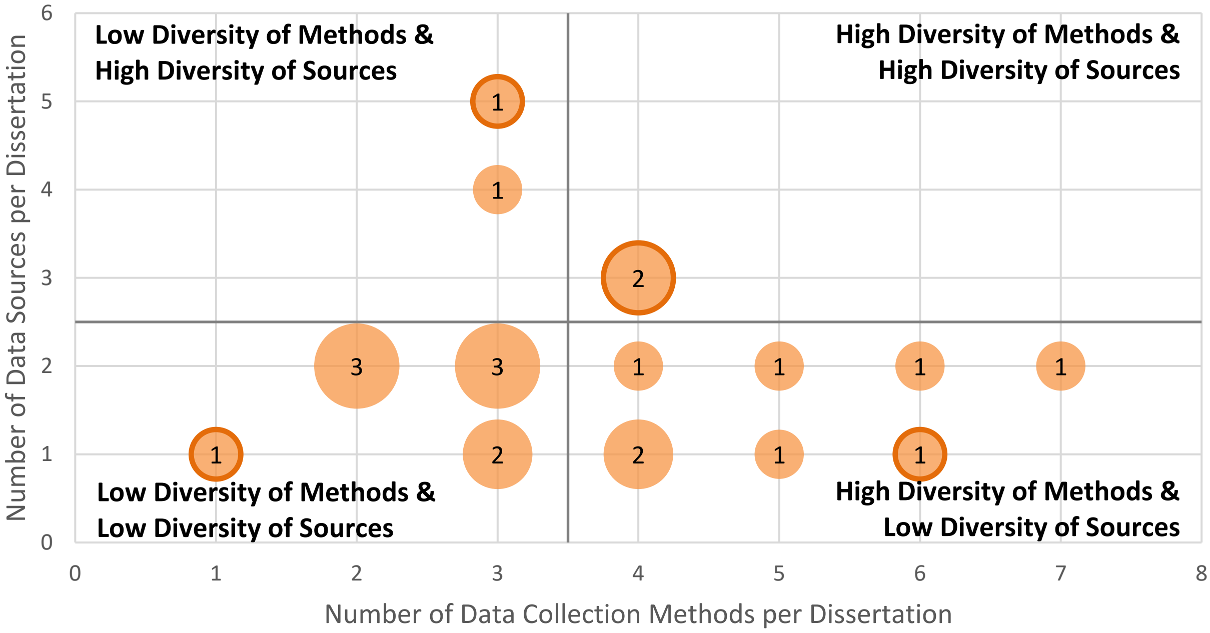


Figure 6. The positioning of dissertations in a research triangulation matrix with two dimensions: data collection methods (x-axis) and data sources (y-axis)

Note: Bubble size is based on the number of dissertations with the same coordinates. Dissertations from the far corner of each quadrant were highlighted.

We further analysed the research triangulation by using a matrix with two dimensions: the number of data collection methods used in each dissertation on the x-axis and the number of data sources used in each dissertation on the y-axis. As Figure 6 shows, the matrix is composed of four quadrants: (1) low diversity of methods and low diversity of sources (lower left quadrant), (2) high diversity of methods and low diversity of sources (lower right quadrant), (3) low diversity of methods and high diversity of sources (upper left quadrant), and (4) high diversity of methods and high diversity of sources (upper right quadrant). The majority of dissertations are located in the lower quadrants; nine dissertations (43%) had low diversity of methods and low diversity of sources, and eight (38%) had high diversity of methods and low diversity of sources. Only two dissertations (Loukomies, 2013; Vartiainen, 2016) had high diversity of methods and high diversity of sources. Table 4 provides four examples of dissertations from the far corner of each quadrant.

Table 4. Four dissertations that illustrate the variation in research methodologies in the dissertations

	Hongisto (2012)	**Ratinen (2016)**	**Tuomisto (2018)**	**Loukomies (2013)**
Collection Methods	Essay	Questionnaire, observations, group interviews, drawing, essays, and lesson plans	Questionnaire, observations, and diaries	Questionnaire, observations, interviews, and meeting memoranda
Data Sources	Students	Pre-service teachers	Teacher educators, pre-service and in-service teachers, students, and peers	Students, teachers, and experts

4.5 Scale, collaboration, and researcher's roles

The scale of the dissertations varied widely in terms of the size of the research team (from an individual researcher to a large multidisciplinary team), the number of research participants (from 15 to over 1000 participants), and the time taken to complete the dissertation (from 3 to 14 years). Eight researchers (38%) conducted their dissertations alone, while the remaining 13 (62%) collaborated with other researchers or disciplines. For example, Ratinen (2016) conducted his dissertation in collaboration with another researcher, and Nousiainen (2008) worked in a multidisciplinary team comprised of members from various fields, including educational sciences, natural sciences, mathematical information technology, game design (e.g., multimedia and graphic design), and stakeholders (e.g., industry representatives, biology and geography teachers, and students from several school levels).

Twenty researchers had one additional role besides that of a researcher. The majority (n = 13, 62%) of the researchers (e.g., Leppäaho, 2007) had three roles: a researcher who plans the research, collects data, and analyses data; a developer who designs and develops a design intervention; and a teacher who teaches in the research intervention. Seven researchers (33%), including Ekonoja (2014), had two roles: a researcher and a developer. Juntunen (2015) was the only one who had a single role: a researcher.

4.6 EDR process

To investigate the EDR processes used by the dissertations, we analysed the phases of EDR, iterations, alternative design interventions, and issues that were considered during the intervention development.

To analyse the EDR phases of the dissertations, we coded the progress of EDR according to three main phases: (1) preliminary research, (2) development phase, and (3) assessment phase (see Plomp, 2013). Although the EDR processes of the dissertations were presented in various ways using various terms (e.g., cases, cycles, phases, stages, and substudies), we found that all dissertations progressed through three main phases. However, the first phase (i.e., investigation of problems, needs, and context) was not fully conducted in several dissertations. For example, Hassinen (2006) did not empirically investigate needs or context and only reviewed the literature on school algebra, curricula, related theories, and textbooks; and Ekonoja's (2014) first phase was conducted as part of his master's thesis. Additionally, while the primary research and assessment phase was reported thoroughly in all dissertations, the development phase was rather brief in some examples (e.g., Oikarinen, 2016) and comprehensive in others (e.g., Juuti, 2005).

As an important characteristic of EDR is its iterative process of design, assessment, and redesign, we investigated the dissertations' iterations by examining revisions of the interventions and the number of multiple subcycles implemented throughout each dissertation (see McKenney & Reeves, 2019). Almost all researchers (n = 20, 95%) revised their interventions during their dissertations. Seven also refined their interventions after their final field trials. With regard to the number of multiple subcycles, 19 researchers (90%) revised their intervention through multiple subcycles. Thirteen (62%) employed two multiple subcycles, four (19%) employed three, one (5%) employed four, and one (5%) employed seven. In addition to performing seven multiple subcycles, Rukajärvi-Saarela (2015) refined her pre- and in-service teacher course after the final field trial. In contrast, two dissertations (10%) performed only one multiple subcycle. After the multiple subcycle, Hassinen (2006) did not revise her Idea-based Algebra teaching model, while Leppäaho (2007) developed his problem-solving materials further in a textbook.

To ensure that their interventions contributed to real-world settings, we also investigated whether any dissertations worked with alternative designs or considered issues besides pedagogy when developing the interventions. No one worked with alternative designs except Nousiainen (2008), whose first project included alternative user interfaces with layouts and different interaction styles and whose second project generated initial ideas and then integrated and developed them in greater detail.

With regard to the issues considered during intervention development, we found that besides pedagogical issues, most of the dissertations considered the needs of policymakers, particularly the National Core Curriculum, when developing interventions. Only a few dissertations considered other issues, such as practicality, usability, administration, and organisation. For example, when developing her ICT learning environment, Aksela (2005) considered pedagogy, the needs of policymakers, practicality (e.g., time, ease of use, resource availability, and classroom space), usability, and technical issues.

4.7 EDR challenges

Finally, we investigated which EDR challenges were encountered during the dissertations. The challenges in the dissertations can be classified into five categories, which are described below.

First, it was difficult to generalise the results due to the small number of research participants, the short length of interventions, the small number of iterative cycles, the insufficiency of relying only on qualitative data, or context-bound research results (e.g., Ekonoja, 2014; Kallunki, 2009). Second, the nature of EDR made it challenging to perform the research for the dissertations. For example, in Nousiainen's (2008) dissertation, it was difficult to compare the research results from different phases, and it was difficult for some participants to recall what happened at the beginning of a long intervention. In the case of Ekonoja (2014), the EDR interventions were typically innovative in nature, and thus there were no previous studies related to his research. Moreover, his intervention relied greatly on technology. Third, the researchers had limited resources in relation to the complexity of EDR, which requires a huge amount of work due to the need to gather and analyse a large dataset (Vartiainen, 2016) and explicitly document the whole process (Pernaa, 2011). Fourth, EDR was often conducted with multidisciplinary collaboration, which required mutual understandings and good teamwork (e.g., Ikävalko, 2017). Fifth, when they took on multiple roles, it was sometimes difficult for the researchers to maintain objectivity (e.g., Oikarinen, 2016).

5 Discussion and conclusions

Our study improves the understanding of how EDR has been utilised and developed and which challenges it has faced over the last two decades by systematically reviewing 21 Finnish doctoral dissertations on mathematics, science, and technology education. The findings indicate that all dissertations made practical and theoretical educational contributions. In line with the literature (e.g., DBRC, 2003; McKenney & Reeves, 2019; Plomp, 2013), all of the dissertations exhibited the characteristics of EDR, including the use of educational problems in practice as a point of departure, research in real-world settings, evolution through an iterative process (i.e., preliminary research, development, and assessment), development of practical interventions, and refining of theoretical knowledge. Moreover, the challenges faced by the researchers (e.g., high demand for conducting EDR with limited resources and the difficulties of multidisciplinary teamwork) are generally similar to those stated by other scholars (e.g., Brown, 1992; McKenney & Reeves, 2019). However, the dissertations were distinctly diverse in terms of the research context (i.e., educational sectors, settings, and domains), educational problems in practice, research outcomes, research methodology (i.e., research methods, data collection methods, and data sources), scale, and collaboration. Like the EDR reviews of Anderson and Shattuck (2012) and Zheng (2015), the findings support the plurality of EDR (see Bell, 2004). Our results indicate that it is feasible to conduct EDR dissertations in different educational sectors, in different settings and domains, at various scales, and with different research designs.

Based on our observations, we agree with other researchers (e.g., Easterday et al., 2017; Ørngreen, 2015; Zheng, 2015) that EDR still needs much more work. Thus, we propose several suggestions for future EDR. First, we encourage agreement between the terms used to describe EDR in different languages to promote consistency and avoid confusion. Second, as EDR is an emergent research approach (Easterday et al., 2017), recent literature should be consulted so that researchers can stay up to date. Third, in agreement with the DBRC (2003) and McKenney and Reeves (2019), we believe that the triangulation of research methods, data collection methods, and data sources is needed to better understand complex authentic phenomena and ensure the trustworthiness of EDR. Fourth, we support Kennedy-Clark (2013) and McKenney and Reeves (2019) and highly encourage multidisciplinary collaboration so that EDR researchers benefit from the expertise of others and increase the feasibility and robustness of their research. Fifth, in line with McKenney and Reeves (2019) and

Ørngreen (2015), when developing the intervention, working with alternative designs and considering various issues faced by all people in real-world contexts can enhance the success of EDR and ensure that the intervention continues to be utilised in real-world settings. Sixth, we agree with McKenney and Reeves (2019), Kennedy-Clark (2015), and Zheng (2015) that design activities and processes should be further emphasised so that others can benefit from them. Finally, due to the appearance of EDR terms in the primary titles of six dissertations, which implies that there is an overemphasis on EDR at the expense of the subject of the research, and the fact that EDR requires substantial resources (Kelly, 2013), we recommend that EDR should be undertaken because of its appropriateness and utility rather than for its own sake.

Our research has several limitations. First, our systematic review included only 21 Finnish dissertations on mathematics, science, and technology education from five universities. A broader dataset in terms of both the number of universities, dissertations, and educational fields would greatly improve the understanding of the utilisation and development of EDR. Second, the large dataset (a total of 4187 pages), the lack of a shared writing structure, and the implicit reporting of information that was necessary for this review made it difficult to perform data coding and analysis. More resources for coding and analysis would increase the precision of the research results and decrease the workload of researchers conducting the review. Last, to gain an overview of the utilisation, development, and challenges of EDR, we adopted a broad perspective when systematically reviewing the use of EDR terms and theoretical frameworks, research contexts, educational problems in practice and research outcomes, research methodologies, the dissertation's scale and collaboration, the researcher's roles, EDR processes, and EDR challenges. While our review indeed provides an overview, a review focusing on specific issues would yield profound insights into EDR.

References

Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher, 41*(1), 16–25. doi:10.3102/0013189X11428813

Bell, P. (2004). On the theoretical breadth of design-based research in education. *Educational Psychologist, 39*(4), 243–253.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences, 2*(2), 141–178.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13.

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches*. Los Angeles: Sage.

Dede, C. (2004). If design-based research is the answer, what is the question? A commentary on Collins, Joseph, and Bielaczyc; diSessa and Cobb; and Fishman, Marx, Blumenthal, Krajcik, and Soloway in the JLS special issue on design-based research. *Journal of the Learning Sciences, 13*(1), 105–114.

Design-Based Research Collective (DBRC). (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 32*(1), 5–8.

Easterday, M. W., Lewis, D. G. R., & Gerber, E. M. (2017). The logic of design research. *Learning: Research and Practice, 4*(2), 131–160. doi:10.1080/23735082.2017.1286367

Edelson, D. C. (2002). Design research: What we learn when we engage in design. *Journal of the Learning Sciences, 11*(1), 105–121.

Juuti, K., & Lavonen, J. (2006). Design-based research in science education: One step towards methodology. *NorDiNa, 4*, 54–68.

Kelly, A. E. (2013). When is design research appropriate? In T. Plomp & N. Nieveen (Eds.), *An introduction to educational design research* (pp. 135–150). Enschede, Netherlands: Netherlands Institute for Curriculum Development.

Kennedy-Clark, S. (2013). Research by design: Design-based research and the higher degree research student. *Journal of Learning Design, 6*(2), 26–32.

Kennedy-Clark, S. (2015). Reflection: Research by design: Design-based research and the higher degree research student. *Journal of Learning Design, 8*(3), 106–107.

McKenney, S., & Reeves, T. C. (2012). *Conducting educational design research. London*: Routledge.

McKenney, S., & Reeves, T. C. (2019). *Conducting educational design research*. New York: Routledge.

McKenney, S., & Reeves, T. C. (2013). Systematic review of design-based research progress: Is a little knowledge a dangerous thing? *Educational Researcher, 42*(2), 97–100. doi:10.3102/0013189X12463781

Official Statistics of Finland. (2015). University education. Appendix table 3. Students in universities and completed university degrees by universities in 2014. Retrieved from http://www.stat.fi/til/yop/2014/yop_2014_2015-05-06_tau_003_en.html

Plomp, T. (2007). Educational design research: An introduction. In T. Plomp & N. Nieveen (Eds.), *An introduction to educational design research* (pp. 9–33). Enschede, Netherlands: Netherlands Institute for Curriculum Development.

Plomp, T. (2013). Educational design research: An introduction. In T. Plomp & N. Nieveen (Eds.), *Educational design research, Part A: An introduction* (pp. 11–50). Enschede, Netherlands: Netherlands Institute for Curriculum Development.

Prediger, S., Gravemeijer, K., & Confrey, J. (2015). Design research with a focus on learning processes: An overview on achievements and challenges. *ZDM Mathematics Education, 47*(6), 877–891. doi: 10.1007/s11858-015-0722-3

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development, 53*(4), pp. 5–23.

Zheng, L. (2015). A systematic literature review of design-based research from 2004 to 2013. *Journal of Computers in Education, 2*(4), 399–420. doi:10.1007/s40692-015-0036-z

Ørngreen, R. (2015). Reflections on design-based research: In online educational and competence development projects. In J. Abdelnour Nocera, B. R. Barricelli, A. Lopes, P. Campos, & T. Clemmensen (Eds.), *Human work interaction design: Analysis and interaction design methods for pervasive and smart workplaces* (pp. 20–38). London: Springer.

List of the 21 selected disserations

Aksela, M. (2005). *Supporting meaningful chemistry learning and higher-order thinking through computer-assisted inquiry: A design research approach* (Doctoral dissertation, University of Helsinki). Retrieved from https://helda.helsinki.fi

Ekonoja, A. (2014). *Oppimateriaalien kehittäminen, hyödyntäminen ja rooli tieto- ja viestintätekniikan opetuksessa* [The development, invocation and role of learning materials in teaching information and communication technology] (Doctoral dissertation, University of Jyväskylä). Retrieved from https://jyx.jyu.fi

Hassinen, S. (2006). *Idealähtöistä koulualgebraa: IDEAA-opetusmallin kehittäminen algebran opetukseen peruskoulun 7. luokalla* [Idea-based school algebra: IDEAA teaching model development for 7th-grade algebra teaching] (Doctoral dissertation, University of Helsinki). Retrieved from https://helda.helsinki.fi

Hongisto, L. (2012). *Luonnontieteen historian opetus tieteenhistorian sisältöjen avulla: Kehittämistutkimus lukion historianopetuksen mahdollisuuksista* [History of natural sciences teaching with the help of history of science content: Design-based research on the possibilities of upper secondary school history teaching] (Doctoral dissertation, University of Helsinki). Retrieved from https://helda.helsinki.fi

Ikävalko, V.-M. (2017). *Mielekkään kemian non-formaalin oppimisympäristön kehittämistutkimus yhteistyössä työelämän kanssa* [Design-based research of a meaningful non-formal chemistry learning environment in cooperation with specialists in the industry] (Doctoral dissertation, University of Helsinki). Retrieved from https://helda.helsinki.fi

Juntunen, M. (2015). *Holistic and inquiry-based education for sustainable development in chemistry* (Doctoral dissertation, University of Helsinki). Retrieved from https://helda.helsinki.fi

Juuti, K. (2005). *Towards primary school physics teaching and learning: Design research approach* (Doctoral dissertation, University of Helsinki). Retrieved from https://helda.helsinki.fi

Kallunki, V. (2009). A historical approach to children's physics education: Modelling of DC-circuit phenomena in a small group (Doctoral dissertation, University of Helsinki). Retrieved from https://helda.helsinki.fi

Leppäaho, H. (2007). *Matemaattisen ongelmanratkaisutaidon opettaminen peruskoulussa: Ongelmanratkaisukurssin kehittäminen ja arviointi* [Teaching mathematical problem-solving skills in Finnish comprehensive school: Design and assessment of a problem-solving course] (Doctoral dissertation, University of Jyväskylä). Retrieved from https://jyx.jyu.fi

Loukomies, A. (2013). *Enhancing students' motivation towards school science with an inquiry-based site visit teaching sequence: A design-based research approach* (Doctoral dissertation, University of Helsinki). Retrieved from https://helda.helsinki.fi

Nieminen, M. (2008). *Ilmavoimien kadetit verkossa: Kokemuksia verkkopohjaisen oppimisympäristön käytöstä matematiikan perusopetuksessa* [Finnish air force cadets in network – experience in use of an online learning environment in basic studies of mathematics] (Doctoral dissertation, University of Jyväskylä). Retrieved from https://jyx.jyu.fi

Nousiainen, T. (2008). *Children's involvement in the design of game-based learning environments* (Doctoral dissertation, University of Jyväskylä). Retrieved from https://jyx.jyu.fi

Oikarinen, J. (2016). *Technology-enhanced statistics learning experiment: A case study at upper secondary level* (Doctoral dissertation, University of Oulu). Retrieved from http://jultika.oulu.fi

Pernaa, J. (2011). *Kehittämistutkimus: Tieto- ja viestintätekniikkaa kemian opetukseen* [Design research: Information and communication technology-based chemistry education] (Doctoral dissertation, University of Helsinki). Retrieved from https://helda.helsinki.fi

Ratinen, I. (2016). *Primary student teachers' climate change conceptualization and implementation on inquiry-based and communicative science teaching: A design* research (Doctoral dissertation, University of Jyväskylä). Retrieved from https://jyx.jyu.fi

Rautiainen, J. (2012). *Kehittämistutkimus: Ongelmalähtöinen kokeellinen korkeakouluopetus* [A design research: Problem- and inquiry-based higher education in chemistry] (Doctoral dissertation, University of Helsinki). Retrieved from https://helda.helsinki.fi

Rukajärvi-Saarela, M. (2015*). Tutkimuksellisuudesta innostusta alakoulujen kemian opetukseen: Kehittämistutkimus osallistavan luokanopettajan perus- ja täydennyskoulutuksen kehittämisestä* [Inquiry-based approach to inspire chemistry education in primary schools. A design research study to develop participatory pre- and in-service teacher education] (Doctoral dissertation, University of Jyväskylä). Retrieved from https://jyx.jyu.fi

Tomperi, P. (2015). *Kehittämistutkimus: Opettajan ammatillisen kehittymisen tutkimusperustainen tukeminen käyttäen SOLO-taksonomiaa - esimerkkinä tutkimuksellinen kokeellinen kemian opetus* [Design research: Research-based support utilising the SOLO taxonomy for teachers' professional development: Inquiry-based chemistry instruction as an example] (Doctoral dissertation, University of Helsinki). Retrieved from https://helda.helsinki.fi

Tuomisto, M. (2018). *Design-based research: Educational chemistry card and board games* (Doctoral dissertation, University of Helsinki). Retrieved from https://helda.helsinki.fi

Vartiainen, J. (2016). *Kehittämistutkimus: Pienten lasten tutkimuksellisen luonnontieteiden opiskelun edistäminen tiedekerho-oppimisympäristössä* [Design research: Promoting natural sciences inquiry-based learning of young children in a science club learning environment] (Doctoral dissertation, University of Helsinki). Retrieved from https://helda.helsinki.fi

Vesterinen, V.-M. (2012). *Nature of science for chemistry education: Design of chemistry teacher education course* (Doctoral dissertation, University of Helsinki). Retrieved from https://helda.helsinki.fi