International Journal of Educational Research 108 (2021) 101782

Contents lists available at ScienceDirect

International Journal of Educational Research

journal homepage: www.elsevier.com/locate/ijedures

Implementing implementation science in a randomized controlled trial in Norwegian early childhood education and care

Ingunn Størksen [a,*], Sigrun K. Ertesvåg [a], Mari Rege [b]

[a] Norwegian Centre for Learning Environment and Behavioural Research in Education, University of Stavanger, Norway
[b] University of Stavanger Business School, University of Stavanger, Norway

ARTICLE INFO

Keywords:
Implementation
Randomized controlled trial (RCT)
Early Childhood Education and Care (ECEC)
Theory of change
Co-production
Implementation framework
Ecological validity

ABSTRACT

The emerging interest in implementation processes in the social, health, and educational sciences has increased the recognition of implementation science. Still, the literature provides limited practical insights on how to successfully develop and conduct interventions within educational settings in randomized controlled trials (RCTs) in line with implementation science. This paper uses the Agder RCT study to provide such insights. We describe how the *theory of change* and *implementation framework* supported systematic tailoring and implementation of a new Early Childhood Education and Care (ECEC) intervention. The paper contributes to the emerging use of implementation science in education in general, and more specifically in ECEC. Finally, we discuss how implementation science also needs to be utilized during the upscaling of the project.

1. Introduction

The call for evidence-based practice continues to grow in education (Gorard, See, & Siddiqui, 2017; Pontoppidan et al., 2018) as well as many other areas of prevention and intervention. Accordingly, several RCT studies to improve Early Childhood Education and Care (ECEC) quality and to improve children's transition to school have been conducted, revealing varying effects (Dyson, Jordan, & Glutting, 2013; Gerholm et al., 2019). One concern is that in many studies the intervention has not been implemented as intended. Meanwhile, the science of putting ideas into action – the science of implementation – has progressed rapidly in all sciences, including the educational sciences (Bertram, Choi, & Elsen, 2018; Domitrovich et al., 2008; Durlak, 2015; Fixsen, Blase, & Van Dyke, 2019; Greenberg, Domitrovich, Graczyk, & Zins, 2005; Humphrey et al., 2016). A growing body of evidence indicates that that the quality of implementation influences desired outcomes (Durlak & DuPre, 2008). Simply put, if we want to achieve outcomes, we must implement with quality. Still, the literature provides limited practical insights and examples of how to develop and tailor an intervention in line with implementation science, and most studies do not provide any evaluation of the implementation of the intervention (Connolly, Keenan, & Urbanska, 2018).

The goal of the study that is presented here, the Agder RCT study, was to develop and scientifically test a new ECEC curriculum to create a better and more even starting point for children before school through practices that combine the Norwegian focus on play with knowledge of learning areas proven foundational for school success. Additionally, the study relied heavily on implementation science. In this paper, we discuss how implementation theory informed the development and implementation of the Agder RCT study

* Corresponding author at: Norwegian Centre for Learning Environment and Behavioural Research in Education, University of Stavanger, 4603 Stavanger, Norway.
E-mail address: ingunn.storksen@uis.no (I. Størksen).

https://doi.org/10.1016/j.ijer.2021.101782
Received 26 January 2021; Received in revised form 31 March 2021; Accepted 1 April 2021
Available online 22 April 2021

(Rege et al., 2019). Importantly, our goal is to describe how 1) the Theory of change and 2) the Implementation framework supported the implementation. The goal of this article is to provide a practical example of how implementation theory, which may seem rather theoretical, can be used to increase intervention implementation quality in ECEC, and strengthen outcomes in RCTs.

1.1. Key concepts from implementation science

Theory of change may be described as a research team's best understanding and hypothesis, based on state-of-the-art scientific knowledge, of what kind of actions and activities may create desired changes (Thornton et al., 2017). In any program or intervention, there should be a rationale for key components and how they are expected to produce change (Bertram, Blase, & Fixsen, 2015; Kellogg Foundation, 2004). Although most RCTs have a theoretical background for change, it is more seldom to see a comprehensive presentation of *how* and *why* change is expected. In this paper, we describe how ECEC research together with implementation science formed the Theory of change in the present case. The study also used co-production to improve user involvement and applicability.

Co-production is a process where citizens and public service professionals initiate, plan, design, and implement new public services (Brandsen, Steen, & Verschuere, 2018). Within the ECEC field, Co-production has the potential to 1) contextualize new programs, and 2) build capacity for playful and responsive teaching among participants. Interestingly, Co-production has many of the same features as playful learning itself, including engagement, meaning-making, iteration, interaction, and joy (Jensen & Morris, 2021). Co-production is expected to lead to measures of change with high *Ecological validity* (Brandsen et al., 2018). Ecological validity refers to whether research findings and new interventions may be generalized and applied to everyday life settings (Wegener & Blankenship, 2019), e.g., how a new curriculum applies to the everyday contexts of ECEC teachers and children.

Furthermore, RCT studies may employ a so-called *Implementation framework* (Bertram et al., 2015) which answers questions of what, how, and who, that are so crucial for successful implementation. An implementation framework can be described as a visual presentation of the main steps and components in a complex implementation process, including a timeframe and milestones.

There are three inter-related goals for the implementation of an intervention: 1) To improve knowledge and understanding of intervention effectiveness, 2) To support implementation, e.g., through the development of research-based frameworks to optimize intervention delivery, and 3) To evaluate the implementation (Humphrey et al., 2016). In this paper, we mainly focus on the two first goals to discuss how implementation theory informed the development and implementation of the intervention. We draw on the perspectives of Fixsen and colleagues (Bertram et al., 2015; Fixsen et al., 2019) that, although not specifically addressing educational settings, provide guidance of high relevance for ECEC settings.

1.2. The Norwegian ECEC context

Norwegian ECEC centers are often considered to be among the best in the world, because of publicly regulated accessibility for children aged one – five years (97 % attendance at age five) and a common Framework Plan for Kindergartens (Ministry of Education & Research., 2017) which focuses on holistic development through care, play, and learning. Free play and outdoor activities – which are seen as crucial to children's well-being and learning - are highly valued and encouraged, and the share of such activities during a typical day is high (Karlsen & Lekhal, 2019; Moser & Martinsen, 2010). However, there are also challenges in Norwegian ECEC, especially when it comes to pedagogical content and learning outcomes (Bjørnestad & Os, 2018; Bjørnestad, Broekhuizen, Os, & Baustad, 2019; Rege, Solli, Størksen, & Votruba, 2018). The Framework Plan only loosely describes learning areas and does not age-differentiate for ages one to five years. This contrasts with research concluding that an age-appropriate curriculum and teachers' knowledge of curriculum is key to high-quality ECEC (Burchinal, 2018; Yoshikawa et al., 2013). There is a gap between the practices and traditions of ECEC centers and schools, and a revision of curricula to ensure pedagogical continuity is recommended (Schleicher, 2019, p. 48). OECD encourages closer collaboration between research and practice in ECEC in Norway (Engel, Barnett, Anders, & Taguma, 2015) to enhance ecological validity. The Agder RCT study addressed these challenges.

2. Theoretical background

2.1. Theory of change

The Theory of Change of the Agder RCT study can be described in line with the Logic Model outlined by the Kellogg Foundation (2004), see Fig. 1. The study started with several *resources (1)*, including ECEC research and theory and implementation theory. A presentation of ECEC research to local funders formed the foundation of the study. Local funders at Agder partnered with researchers with the overall goal to improve school readiness skills and school success for all children in their district and local teachers were soon involved as partners. These resources were assumed to form a solid base for the main *activities (2)* in the project, including teacher training and coaching, co-production, implementation of a new playful learning curriculum, and the implementation of a large RCT study focusing on children's school readiness skills. This was assumed to give *outputs (3)* at the teacher level in terms of enhanced teacher skills, teacher satisfaction, and teacher motivation, and in terms of a new playful learning curriculum. Jointly, these outputs were assumed to result in important *outcomes (4)* at the child level. We expected strengthened school readiness skills among study children at the end of the intervention year, at one-year follow-up, and – in the future – mid-term and long-term effects. Finally, we assume that the Agder RCT study will continue to *impact society (5)*, through the dissemination of research results and the new playful learning curriculum, which will, in turn, give children better school readiness skills.

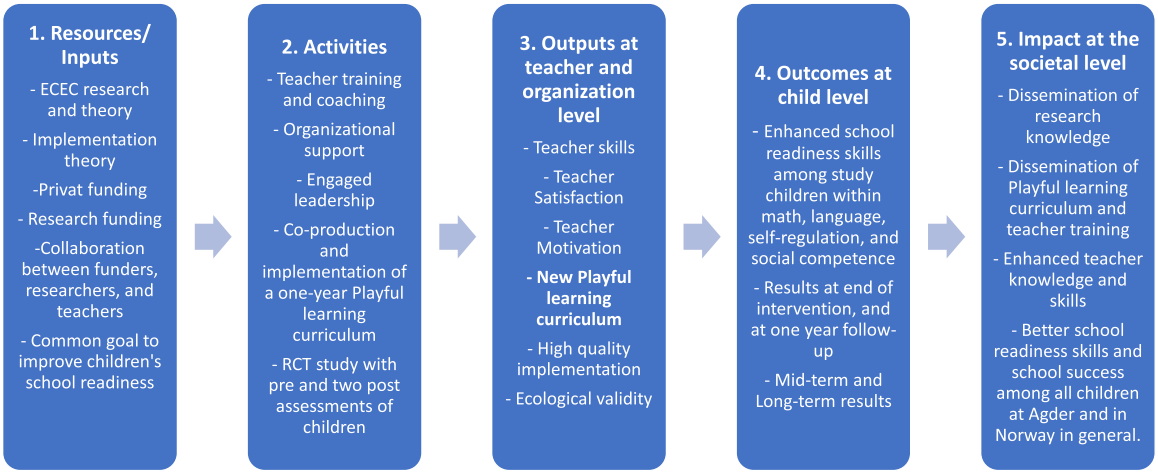

Fig. 1. The theory of change for the Agder RCT study presented in line with the Logic Model from the Kellogg Foundation.

The theoretical foundation and core component of the Agder RCT study are outlined in more detail in Fig. 2. Importantly, it included critical insights and core components from both the ECEC field and from Implementation Science. Identifying core components of an intervention is a key to successful implementation (Bertram et al., 2015). During the first phase of the Agder RCT study, the researcher reviewed relevant research and theory, see dark blue boxes. This helped identify the core components of the intervention and implementation, represented in the two largest boxes. Together these core components were expected to create a high-quality implementation and to facilitate positive outputs for teachers, important outcomes for children, and high ecological validity of the new curriculum intervention, as visualized by the two boxes to the right.

2.2. *Core components of the ECEC intervention*

Early learning areas identified as foundational for young children's school success include early mathematics, language and literacy, self-regulation, and social competence. Certain pedagogical approaches are identified in research as essential in the teaching of young children in transition to school, including supportive relationships and playful learning approaches, see Fig. 2.

Supportive relationships between teachers and students are crucial for young children's well-being and learning. Positive relationships with children in general and especially with children at risk can reduce the effect of early adversities and promote healthy social and academic child development (Pianta, 1999). High-quality interactions in pre-K are found to be related to children's learning outcomes (Sabol, Soliday Hong, Pianta, & Burchinal, 2013). Supportive relationships between teachers and children in the Agder RCT study were encouraged by strengthening teacher skills in this domain. This was achieved through training, practical and written

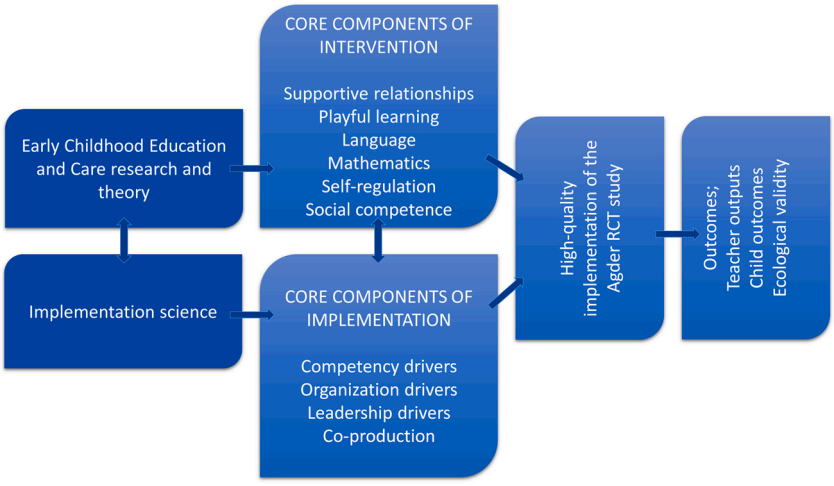

Fig. 2. The theoretical foundation in the Agder RCT study includes Core Components of the ECEC Intervention and Core Components of the Implementation.

assignments, and coaching.

Playful Learning research and theory show that children's early learning is most efficient during play and discovery together with other children and with teachers (Fisher, Hirsh-Pasek, Golinkoff, Dinger, & Berk, 2011; Hirsh-Pasek, Michnick Golinkoff, Berk, & Singer, 2008). How you learn is as important as what you learn. Learning takes place during play, both during *free play* and *guided play* where teachers intentionally prepare and introduce books, games, activities, or toys, and engage children in exploring core themes (Weisberg, Hirsh-Pasek, & Golinkoff, 2013). Children are active, engaged, interact with others, and experience the activity as meaningful (Hirsh-Pasek et al., 2008). In the present case, playful learning skills and pedagogical approaches were strengthened among teachers through training, group discussions, practical and written assignments, and coaching.

Children need *Social competence* to handle the complex social contexts of school. Learning-related skills such as self-regulation and social competence contribute to early school success (McClelland, Acock, & Morrison, 2006). Early *self-regulation* is foundational to succeed in school. Preschool behavioral regulation has been linked to emergent literacy, vocabulary, and math skills (Lenes, McClelland, ten Braak, Idsøe, & Størksen, 2020; McClelland et al., 2007; ten Braak, Størksen, Idsoe, & McClelland, 2019). In the present case, teachers were trained in the importance of early social competence and self-regulation and practiced and received coaching on activities and games for children. At the child level, age-appropriate games to improve children's social competence and self-regulation skills were introduced.

Language and literacy skills during early childhood form the basis for future learning (Duncan et al., 2007). Studies show that interactive book reading in combination with professional development for teachers is effective in promoting richer classroom conversations and gains in child vocabulary and oral comprehension (Bierman et al., 2008). Exploring rimes, letters, and sounds at an early age is generally seen as an important pre-reading activity (Muter, Hulme, Snowling, & Stevenson, 2004). In line with this, teachers received training and coaching, and administered games for children that stimulated skills such as phonological awareness and vocabulary.

The importance of early *Mathematics* is increasingly evident, and the predictive power of early math on later math skills is considerable (ten Braak & Størksen, 2021). Research shows that the strongest early childhood predictors of general school achievement include school-entry math, reading, and attention skills, and math skills stand out as the most important predictors (Claessens & Engel, 2013; Duncan et al., 2007). Teachers in the Agder RCT study expressed a special need for input in this domain through a pre-intervention assessment, and training was intensified to strengthen teacher skills (Rege et al., 2019). Games and activities were created to stimulate child skills such as numeracy, measurement, geometry, and statistics (Størksen et al., 2018).

2.3. Core components of the implementation

Implementation science can help implement interventions with fidelity by identifying specific components of the implementation process critical to change. In this study, we were inspired by the theory from Bertram on implementation drivers (Bertram et al., 2015, 2018), including *competency drivers, organization drivers,* and *leadership drivers* which are building blocks of the infrastructure needed to support organizational change (Fixsen et al., 2019). Each implementation driver includes several sub-components.

Competency drivers include 1) selecting participants, followed by 2) high-quality training, and coaching, and 3) fidelity testing. Selecting competent staff within the specific field of innovation is often mentioned as the first success criteria in implementation theory, and the purpose of competency drivers is to facilitate new learning, competence, and confidence through training and coaching. Thus, training and coaching of competent staff are the primary competence drivers in successful implementation (Bertram et al., 2015). During the initiation of the Agder RCT study, all centers in this district were invited to participate. Because of the importance of selecting qualified staff, we chose to focus on qualified teachers, and thus a requirement from the project was that the lead teacher(s) of the preschool group(s) had a college education in ECEC and volunteered to participate. A total of 71 centers signed up for the project, each with one or two voluntary teachers. This self-selection led to the recruitment of highly qualified and motivated staff.

The teacher training consisted of high-quality teaching within core components of the intervention from the research team. In line with implementation science, participant involvement was held high. The teacher training included both improvement cycle loops, dialogue, and coaching. Teacher involvement in training through co-production was encouraged to strengthen ownership and anchoring of the project. For example, during the first teacher training session, the initial activity was letting teachers list up characteristics they would like to see in the new curriculum. This led to meaningful discussions and facilitated collaboration. Both the teacher training year and the implementation year contained coaching from the research team. During training, coaching was given orally in plenary and groups, and individually as feedback to written assignments. During the implementation, we arranged two workshops and several individual phone meetings. New solutions were collaboratively discovered.

Several activities to strengthen and assess fidelity were chosen. Teachers handed in papers as part of the teacher training in which they integrated theory and practice, and they were subsequently given feedback and coaching. All participants passed the 15-credit point exam at the end of the teacher training. Participating centers signed an agreement to spend at least eight hours per week engaging children in the new curriculum. Fidelity was assessed through electronic weekly assessments asking teachers how many hours were spent on activities and games during the previous week. Among the reports that were submitted, 67 percent reported spending eight or more hours on the learning activities the previous week, and only in 16 percent of the reports, the teacher reported they spent less than six hours on games and activities with the preschool group (Rege et al., 2019) which was considered to be satisfying results. A qualitative observation study of the implementation in the mathematical domain investigated fidelity in terms of playful learning processes in participating centers. Activities gave opportunities for children's mathematical learning (Hundeland, Carlsen, & Erfjord, 2020).

Organization drivers consist of 1) systems intervention, 2) a facilitative administration, and 3) decision support data systems (Bertram et al., 2015; Fixsen et al., 2019). Systems intervention relates to the way the organization resolves practical issues related to implementation, e.g., work with external funders to ensure enough resources to support staff during implementation. In this project, participating centers were supported with resources for substitute teachers during training and development. This support was organized through the center leaders. According to implementation science, facilitative administrators look for ways to support practitioners' use of an innovation. They work to achieve and sustain high fidelity use of innovation, and data is used to support decision making (Bertram et al., 2015; Fixsen et al., 2019). In the initiation of the Agder RCT study, ECEC center leaders and teachers were required to commit to facilitating implementation by signing a contract. This rigorous procedure probably strengthened implementation and reduced attrition at the center level. In fact, during this five-year-long project, only one center withdrew from participation. Throughout the project, all data collected through oral and written feedback from teachers continuously informed the researchers on the decision-making and counseling of participants, and thus they constituted a decision support data system in the project.

Leadership drivers (Fixsen et al., 2019) imply that leaders must fully agree to engage in the implementation process. Their role is to create supportive learning environments at all levels of the organization. Leadership is a complex process, especially when many parties are involved that have never worked together before. Unresolved issues and uncertainties will make this process even more complex. Leadership drivers consist of 1) Technical leadership, which is required when the involved parties agree upon problems and solutions, and 2) Adaptive leadership, which is required when new challenges arise, or when the definition of problems and their solutions are less clear (Bertram et al., 2015). In the Agder RCT study, technical and adaptive leadership were required both at the research team level and at the ECEC center level.

The leaders of the Agder RCT study organized and led the project according to the implementation framework (see Fig. 4). This part of the leadership may be defined as technical, related to conducting activities as planned. However, the research project also required adaptive leadership to handle unforeseen challenges. For example, recruitment to the project required more attention than planned, and regional information meetings were set up in several locations in the district of Agder to inform centers about project enrollment.

Center-level implementation of this project also required both technical and adaptive leadership. On the technical side, the center leaders signed a contract describing what was expected from them and conducted activities according to predefined plans. Adaptive leadership was required when unexpected challenges arose, e.g. sick leaves among staff or requirements for rooms, locations, or materials for certain activities. Bertram et al. (2015) describe how the various implementation drivers are compensatory because weaknesses in one driver can mitigate by strengths in another driver. In the Agder RCT study, we built up very strong competency drivers, and during the implementation of the intervention center leaders and teachers oftentimes found their own solutions to challenges.

2.4. Co-production

Translating and implementing research-based knowledge into practice in ECEC-settings is a complex process. Prescriptive practices where researchers script a very detailed intervention for teachers may contradict the very idea of playful learning and responsive teaching. Therefore co-production is recommended both during the contextualization of programs and as a means to build capacity for playful and responsive teaching among teachers (Jensen & Morris, 2021).

To develop an ecologically valid innovation, the project team organized a co-production process between teachers and researchers. High ecological validity leads to effective innovation in the sense that the intervention can produce change and fit well with the practice field (Wegener & Blankenship, 2019). The goal for the new curriculum was not only that it should be based on research, but also that it should align well with traditions and values in the Norwegian Framework Plan (Ministry of Education & Research., 2017). Therefore, during the development of the new curriculum, principles of co-production and rapid cycle iterative learning (Brandsen et al., 2018; Osborne & Strokosch, 2013; Schindler, Fisher, & Shonkoff, 2017) between teachers and researchers were applied. In *rapid cycle iterative learning* (Schindler et al., 2017) adjustments and improvements can be made to strengthen a new intervention through series of iterative steps.

Teachers in the Agder RCT study were spread across a large geographical area, and to reduce travel days, the lectures were organized in four two-day workshops. During every workshop, lectures, group work, and plenary discussions related to the six core elements of the intervention took place. Additionally, implementation theory was presented so that the teachers themselves could utilize this theory in their local implementations. At the end of each workshop, drafts and ideas for activities and games that could

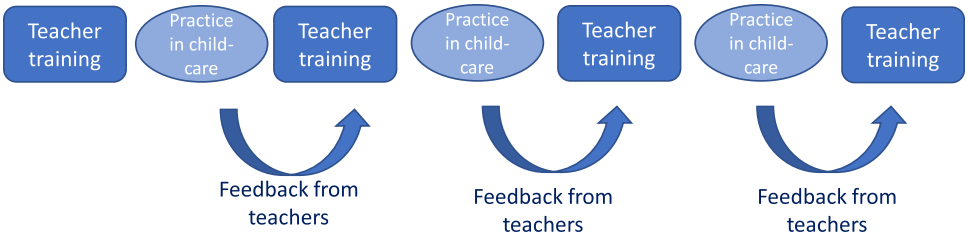

Fig. 3. The iterative co-production process in the Agder RCT study.

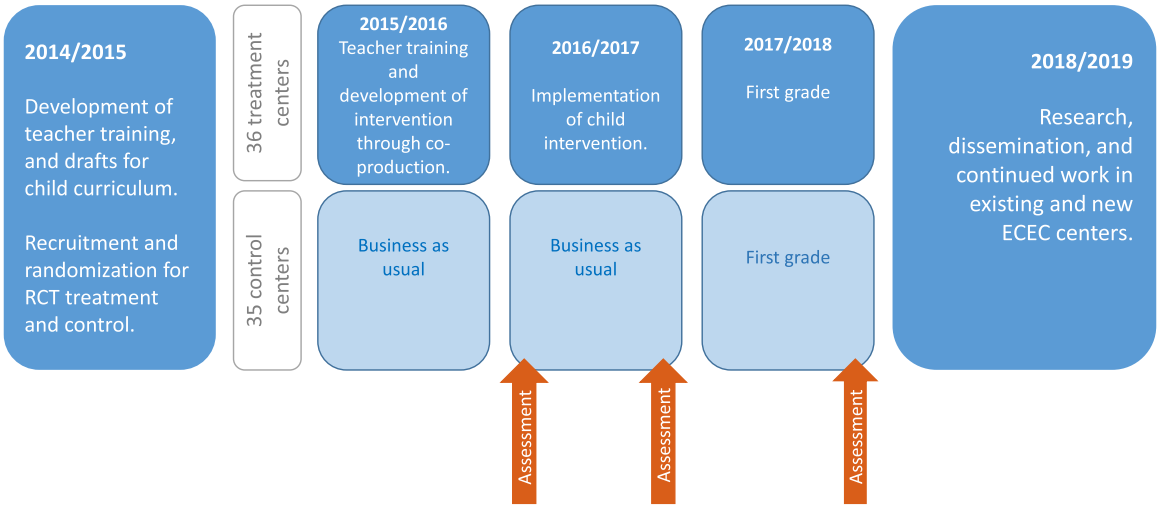

Fig. 4. Implementation framework in the Agder RCT study.

potentially augment child development within the core themes of the intervention were presented. Teachers brought the activities and games back to their local centers and tried them out with children. At the next workshop, they gave oral and written feedback and reflections to researchers, and games and activities were adjusted or replaces with better alternatives. Furthermore, challenges and successes related to pedagogical practices and plans were discussed, e.g., teachers quite immediately experienced that playful learning was most feasible in smaller groups compared to large circle time groups. New ideas were integrated into the curriculum intervention, and in this way, rapid cycle iterative learning and co-production took place, see Fig. 3.

The co-production resulted in 130 games and activities that were summed up in the new semi-structured curriculum intervention. Elaboration and refinement of pedagogical approaches were also part of this co-production cycle. Discussions between teachers and researchers resulted in meaningful reflections and collaborative learning, that was brought into the new curriculum intervention. The resulting curriculum intervention contained a presentation of the core components, associated games and activities, pedagogical approaches, monthly progression plans, and templates for planning days and weeks. The new curriculum was printed in a book (Størksen et al., 2018) that the publisher agreed to hold back from public release until after the research project had ended. Thus, the curriculum was exclusive for the intervention group during the implementation period which is highly recommended in RCT studies.

3. Research design

The research design consisted of a classical set-up for clustered randomized controlled trials. Thorough details of the research design and results of the Agder RCT study are described elsewhere (Rege et al., 2019). Although the purpose of the present study is to exemplify the value of utilizing implementation theory in RCT studies, to enhance understanding of the present case, the overall research design will now be presented.

3.1. Implementation framework

The theory of change outlined above formed the theoretical expectations for the Agder RCT study. In addition, we created a more specific Implementation Framework identifying crucial steps and a timeframe for central activities to be carried throughout the project, see Fig. 4. We planned to spend our first year (2014–2015) developing teacher training and drafting potential activities for the curriculum intervention for children. Furthermore, we expected to spend time on the recruitment of ECEC centers, teachers, and children to manage to reach a satisfactory sample for our RCT. During the next year (2015–2016), we planned to give teachers in the intervention group their training and to co-produce the curriculum intervention together with these teachers. The following year (2016–2017), was dedicated to the implementation of the curriculum intervention with the children, and to pre- and post-assessment. After first grade, children were to be assessed again for follow-up data.

With a large number of centers, children, and teachers, and 13 members in the research team, this implementation framework helped organize small activities (e.g. information meetings or consent forms) and larger activities (e.g. 3 data-collections with 701 children at each timepoint) in a sequenced way to support the successful implementation of the curriculum intervention as well as the overall RCT. Implementation of the child intervention was supported by the semi-structured ECEC curriculum that had been developed by teachers and researchers (Størksen et al., 2018) together with the training and coaching of teachers.

The actual implementation of the Agder RCT study followed the implementation framework closely, although adjustments had to be made at some stages. Still, having an overall implementation framework enabled us to realize the consequences of adjustments and changes for future parts of the project. During teacher training, teachers and researchers collaborated in the co-production of the curriculum intervention, and, in the next school year, teachers implemented the new curriculum with their five-year-old children. All

children were assessed pre- and post-intervention, and at one-year-follow-up during first grade.

Centers in the intervention group (the treatment group) had agreed to work specifically with activities from the curriculum for 8 h per week. The curriculum was designed to promote what is called guided play within the Playful Learning theory (Hirsh-Pasek et al., 2008). Teachers created settings where children were active, engaged, and interacted with teachers and peers in meaningful activities. The curriculum included plans and activities for each month, yet teachers were also strongly encouraged to make local adjustments during implementation related to e.g. local facilities or specific needs or interests in each child group (Størksen et al., 2018). Thus, we sought a balance between intervention fidelity and local relevance as recommended in contemporary ECEC implementation literature (Jensen & Morris, 2021).

3.2. Sample

Information meetings for all municipalities and ECEC centers in the Agder region were set up for recruitment to the Agder RCT study. Out of 30 municipalities in the region, 15 signed up for the project. Within these municipalities, ECEC center leaders and their teachers signed up voluntarily through written contracts. Among the 190 ECEC centers in these municipalities, 72 signed up for the study. However, before the intervention year, one ECEC center in the control group withdrew, leaving 71 participating centers. Centers were randomized to the intervention and control groups using block randomization based on location (municipality) and center size. Parental consent for the five-year-old children was collected before randomization (n = 569). However, due to a large timeframe between consent collection and the intervention year, some parents signed up after the invitation had closed and before the intervention had started (n = 132) giving a total sample of 701 children. Late consent was more prevalent in the intervention group. Since participants who sign up late may have specific characteristics, late consent was controlled for in the analyses (Rege et al., 2019).

4. Results

The Agder RCT study tested a comprehensive ECEC curriculum targeting four learning areas recommended as crucial for children's future learning and adjustment in the universal play-based context of Norway. The main outcomes from the study are reported elsewhere (Hundeland, Carlsen, & Erfjord, 2020; Rege et al., 2019) but will also be outlined here to illustrate how implementation theory may strengthen intervention outcomes in RCT research.

4.1. Teacher outputs

At the end of the intervention year, we sent out electronic evaluation assessments to the teachers. A total of 38 out of 42 teachers (90.5 %) in the treatment centers replied to this assessment. In general, the teachers were very satisfied with the training. On a question asking for their overall evaluation of the training, 88 % described the training as "very good", 12 % reported that it was "good", and no teachers responded "medium", "weak" or "very weak". For more teacher results, see Table 1.

Teachers were asked whether their working methods in the groups with five-year-olds had changed or whether they were similar to previous practices. Since there is no prevailing curriculum for five-year-olds in Norway, many centers arrange weekly meetings with this group to give them some preschool training through activities quite like first-grade activities. As can be seen in Table 2, teacher replies indicate a trend towards more playful learning activities and fewer desk activities for the children during the intervention year, although some teachers reported that they worked in line with previous practices. A few teachers conducted more desk (14.3 %) and worksheet (14.3 %) activities for the children during the intervention year. The varying results may reflect the large variety in pedagogical approaches before the intervention (Rege, Solli, Størksen, & Votruba, 2018), and this may be related to very loosely described expectations for teachers in the Framework Plan (Ministry of Education & Research., 2017). Another possibility is that the few teachers with negative responses to this evaluation oppose the pedagogical approaches of the Agder RCT study. The Norwegian ECEC tradition highly values free play and outdoor activities (Karlsen & Lekhal, 2019; Moser & Martinsen, 2010) and the Framework Plan contains no benchmarks for children's learning. Many Norwegian ECEC teachers strongly uphold the intrinsic value of childhood and free play, and this value is also stated in the Framework Plan. Therefore, a few of the teachers may have experienced the changes caused by the study (eight hours of guided play per week) as negative compared to the prevailing traditions and expectations.

Although only 38 out of 42 teachers in the intervention group replied to this evaluation, the results align with our theory of change

Table 1
Teachers' replies to statements regarding activities.

How much do you agree with the following statement?	Disagree strongly	Disagree	Disagree slightly	Agree slightly	Agree	Agree Strongly
Working with the activities in the Agder project has been meaningful				3.6 %	28.6 %	67.9 %
The children have enjoyed the activities in the Agder project					35.7 %	64.3 %
The children have learned a lot from the activities in the Agder project					28.6 %	71.4 %
In my center, we will continue to work with activities from the Agder project during the forthcoming 5-year period			3.6 %	7.1 %	50 %	39.3 %

Table 2
Teachers' replies to questions related to work methods.

	More than previously	The same amount as previously	Less than previously
To what extent did the children use pre-printed worksheets or workbooks in the group for five-year-olds during the past year?	14.3 %	17.9 %	67.9 %
To what extent have the children in the group of five-year-olds worked at desks or tables during the past year?	14.3 %	28.6 %	57.1 %
To what extent have you worked with playful learning activities in the group for five-year-olds during the past year?	92.9 %	7.1 %	

and give preliminary support to the notion that the Agder RCT study led to positive teacher experiences and new playful learning practices for a majority of the teachers. This was expected to lead to positive child outcomes.

4.2. Child outcomes

Early childhood mathematics skills were assessed with the *Ani Banani Math Test* (ABMT; Størksen & Mosvold, 2013; ten Braak and Størksen, 2021). The language and literacy skills were assessed with the *Norwegian Vocabulary Test* (NVT; Størksen, Ellingsen, Tvedt, & Idsøe, 2013) and a 12-item *blending task* that is part of the official literacy screening battery from The Norwegian Directorate for Education and Training. Three assessments that measured early self-regulation were included: The *Head-Toes-Knees-Shoulders* task (HTKS; McClelland et al., 2014), the *Hearts and Flowers* task (Davidson, Amso, Anderson, & Diamond, 2006), the *Forward/Backward Digit Span* subtest from the Wechsler Intelligence Scales for children-III (Wechsler, 2004). Mean scores were calculated within each domain and standardized.

A variable for treatment status (intervention versus control) was entered in a regression analysis with post-test assessment scores for children as outcomes, controlling for baseline test scores, indicators for randomization block, gender, birth month, parental characteristics (mother and father's education, earnings, and an indicator country of birth). Analyses were clustered at the center level to adjust for correlated error terms within centers.

Significant treatment effects were found for a sum score of all assessments (ES = .10), for math (ES = .13), and for self-regulation (ES = .11), but no significant effects were found in the language and literacy domain. A similar regression was conducted with one-year-follow-up assessments (spring of first grade) almost one year after the intervention. Significant treatment effects were found for a sum score of all assessments (ES = .13) and for math (ES = .23), but no significant treatment effects were found for self-regulation or the language and literacy domain. Sub-group analyses revealed that treatment effects were almost entirely driven by effects in centers with low pre-intervention quality (Rege et al., 2019).

A qualitative follow-up study focused on the mathematical discourses in participating ECEC centers. Systematic observations revealed that the extent and nature of verbal participation among children and the children's mathematical engagement were higher in treatment centers than in the control centers (Hundeland, Carlsen, & Erfjord, 2020), and thus the quantitative and qualitative findings in the Agder RCT support each other. Results at the child level support our expectations outlined in the theory of change. We expected that theory-based teacher training and coaching, co-production, and implementation of a new playful learning curriculum, would give important outputs in terms of playful learning professes led by teachers that again would result in important outcomes at the child level.

4.3. Ecological validity

Based on the theory of change, we facilitated co-production and rapid cycle iterative learning between teachers and researchers to achieve high ecological validity (Wegener & Blankenship, 2019). Thus, the curriculum was intentionally constructed to match well with traditions and values in the Norwegian ECEC. During the co-production process, teachers handed in written assignments with reactions, comments, and suggestions for the playful learning activities and games for the intervention. They agreed to print some of their comments in the intervention book (Størksen et al., 2018), and some of these are cited here to demonstrate experiences during the co-production phase.

Teachers tested several ways of doing an interactive book reading with children during the co-production and rapid cycle iterative learning. Although they had been reading with children before, this way of reading seemed to be an eye-opener and a positive experience; an element they wanted to develop and integrate as part of the intervention. Teachers chose books for their local child groups and made various experiences:

> *Since I was confident of the content of the book and had thought carefully through focal words and the content of the pictures, I was able to communicate freely with the children while simultaneously keeping track of the storyline* (Størksen et al., 2018).
> *The children showed joy in the child group and played out the characters from the book we had read* (Størksen et al., 2018).
> *It was a very exciting and intense reading session. We all had a good experience and all children participated and were part of creating something together.* (Størksen et al., 2018).

One teacher reported multiple possibilities for applying a self-regulation game he tested out in his child group:

> *Children love games that involve running and competition. It was easy to engage the children in this activity. This was a fun game that may well be organized both inside and outdoors, spontaneously, or organized, either with or without adults* (Størksen et al., 2018).

Several teachers suggested new games and activities, e.g., games related to social competence and collaboration:

> *I suggest we include "the reverse chair game" in the preschool intervention. One chair is removed every time (the music stops), but no children leave the game. The task is that the same group of children must find space on fewer chairs. They need to collaborate and share the space* (Størksen et al., 2018).

The teachers and children showed a preference for math games that involved play and physical activity, for example throwing paper airplanes and measuring the length of their throw:

> *They are very interested in flying with paper airplanes and think it is fun. This (the game) led to a play-based approach to practicing what a meter is, which many of the children had little knowledge of from before* (Størksen et al., 2018).

More comments and feedback from teachers can be read in the book containing the curriculum (Størksen et al., 2018). Note that the reported quotes in the book relate to the activities that were finally selected for the curriculum, and therefore there is a bias towards positive responses. Games or activities for children that received negative teacher evaluations were adjusted or replaced by other activities through iterations and adjustments in the co-production process. The curriculum intervention was created by the teachers themselves in co-production together with researchers, and thus it is not surprising that they gave good evaluations. In general, the teacher reactions support the notion that we achieved ecological validity in the final version of the curriculum intervention through the active use of co-production as a part of our theory of change. Future research and experience will teach us more about how the new curriculum intervention will generalize and apply to everyday ECEC settings in Norway in general.

5. Impact at a societal level

The purpose of the Agder RCT study was to investigate whether a research-based, and age-appropriate ECEC curriculum would give children a better starting point for school compared to business as usual. To strengthen fidelity and outcomes, we had a high focus on implementation, e.g., through co-production, and self-selection of motivated teachers, in addition to generous support for teachers through substitute teachers, extensive training, practice, and coaching. However, the trade-off of this procedure is that this very costly high-quality implementation may have driven the results. The question remains if it is possible to scale up the intervention to a wider population of ECEC centers. Evaluating upscaling of educational interventions to the district level has been highlighted as important in previous research (Tymms et al., 2011).

A crucial next step is to study external validity, by applying implementation science in a more cost-efficient way during the upscaling of the new ECEC curriculum. The project has resulted in a curriculum book, a related webpage with lectures and examples of games and activities (www.lekbasert.no), and a partly web-based professional development (PD) course for ECEC teachers. A follow-up RCT will investigate whether the book, the webpage and a one-day PD training for teachers constitute enough implementation support for ECEC centers to achieve positive school readiness outcomes for children. This will give important insight into how implementation theory can be applied when scaling up interventions that have been co-produced. We will inform the teachers about the co-production process and the findings from the Agder RCT study to establishing ownership and engagement. This is expected to create ownership since co-production occurred within their own profession and cultural context, and engagement since the curriculum has proven effective on important child outcomes.

6. Summary

In the Agder RCT study, a well-founded theory of change, and a detailed implementation framework was created. As described, there is reason to believe that state-of-the-art implementation science together with clearly defined core components from the ECEC field supported high-quality implementation. Our review of outcomes support assumptions in our Theory of change and indicate that implementation was successful at the child and teacher level and that ecological validity was achieved. The case lends support for the value and advantages of planning RCTs through a comprehensive theory of change and implementation framework. Findings encourage future ECEC projects to utilize implementation science in research and everyday practices.

Funding

The Agder RCT study was funded by the Research Council of Norway through grants no. 237973 and 270703, The Sørlandet Knowledge Foundations, and The Agder County administrations.

Ethics approval

The Agder RCT study has been approved by the Norwegian Centre for Research Data and the Norwegian Data Inspectorate.

Informed consent

All participants in the study gave informed consent. Child participants have parental consent to participate.

Declaration of Competing Interest

There are no conflicts of interest to report.

References

Bertram, R., Blase, K., & Fixsen, D. L. (2015). Improving programs and outcomes: Implementation frameworks and organization change, 01/01 *Research on Social Work Practice, 25*, 477–487. https://doi.org/10.1177/1049731514537687.

Bertram, R., Choi, S.-W., & Elsen, M. (2018). Integrating implementation science and evidence-based practice into academic and field curricula, 06/07 *Journal of Social Work Education*, 1–11. https://doi.org/10.1080/10437797.2018.1434441.

Bierman, K. L., Domitrovich, C. E., Nix, R. L., Gest, S. D., Welsh, J. A., Greenberg, M. T., et al. (2008). Promoting academic and social-emotional school readiness: The head start REDI program. Nov-Dec *Child Development, 79*(6), 1802–1817. https://doi.org/10.1111/j.1467-8624.2008.01227.x.

Bjørnestad, E., & Os, E. (2018). Quality in norwegian childcare for toddlers using ITERS-R. European early childhood education. *European Early Childhood Education Reach Journal, 26*(1).

Bjørnestad, E., Broekhuizen, M. L., Os, E., & Baustad, A. G. (2019). Interaction quality in norwegian ECEC for toddlers measured with the caregiver interaction profile (CIP) scales. *Scandinavian Journal of Educational Research*, 1–20. https://doi.org/10.1080/00313831.2019.1639813.

Brandsen, T., Steen, T., & Verschuere, B. (2018). *Co-production and Co-creation : Engaging citizens in public services*. Taylor & Francis.

Burchinal, M. (2018). Measuring early care and education quality. *Child Development Perspectives, 12*(1), 3–9. https://doi.org/10.1111/cdep.12260.

Claessens, A., & Engel, M. (2013). How important is where you start? Early mathematics knowledge and later school success. *Teachers College Record, 115*.

Connolly, P., Keenan, C., & Urbanska, K. (2018). The trials of evidence-based practice in education: A systematic review of randomised controlled trials in education research 1980–2016, 2018/07/03 *Educational Research, 60*(3), 276–291. https://doi.org/10.1080/00131881.2018.1493353.

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. // *Neuropsychologia, 44*(11), 2037–2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006.

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., et al. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: A conceptual framework, 2008/01/01 *Advances in School Mental Health Promotion, 1*(3), 6–28. https://doi.org/10.1080/1754730X.2008.9715730.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., et al. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428–1446.

Durlak, J. A. (2015). What everyone should know about implementation. *Handbook of social and emotional learning: Research and practice* (pp. 395–405). The Guilford Press.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*(3–4), 327. https://doi.org/10.1007/s10464-008-9165-0.

Dyson, N. I., Jordan, N. C., & Glutting, J. (2013). A number sense intervention for low-income kindergartners at risk for mathematics difficulties. March/April 2013 *Journal of Learning Disabilities, 46*(2), 166–181. https://doi.org/10.1177/0022219411410233.

Engel, A., Barnett, S., Anders, Y., & Taguma, M. (2015). *Early childhood education and care policy review NORWAY*. https://www.regjeringen.no/contentassets/ee68f4c948b64c1fbc45baa91f288472/oecd-norway-ecec-review_final_web.pdf.

Fisher, K. R., Hirsh-Pasek, K., Golinkoff, R. M., Dinger, D. G., & Berk, L. E. (2011). Playing around in school: Implications for learning and educational policy. In A. D. Pellegrini (Ed.), *The Oxford handbook of the development of play* (pp. 341–360). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195393002.001.0001.

Fixsen, D. L., Blase, K. A., & Van Dyke, M. K. (2019). Implementation practice and science. *The active implementation research network*.

Gerholm, T., Kallioinen, P., Tonér, S., Frankenberg, S., Kjällander, S., Palmer, A., et al. (2019). A randomized controlled trial to examine the effect of two teaching methods on preschool children's language and communication, executive functions, socioemotional comprehension, and early math skills [journal article]. September 05 *BMC Psychology, 7*(1). https://doi.org/10.1186/s40359-019-0325-9.

Gorard, S., See, B., & Siddiqui, N. (2017). The trials of evidence-based education: The promises. *Opportunities and Problems of Trials in Education*. https://doi.org/10.4324/9781315456898.

Greenberg, M., Domitrovich, C., Graczyk, P., & Zins, J. (2005). *The study of implementation in school-based preventive interventions: Theory, research, and practice*, 01/01 (p. 3).

Hirsh-Pasek, K., Michnick Golinkoff, R., Berk, L. E., & Singer, D. (2008). *A mandate for playful learning in preschool: Presenting the evidence*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195382716.001.0001.

Humphrey, N., Lendrum, A., Ashworth, E., Frearson, K., Buck, R., & Kerr, K. (2016). *Implementation and process evaluation (IPE) for interventions in educational settings: An introductory handbook*.

Hundeland, P. S., Carlsen, M., & Erfjord, I. (2020). Qualities of mathematical discourses in kindergartens. *ZDM, 524*, 691–702. https://doi.org/10.1007/s11858-020-01146-w.

Jensen, H., & Morris, A. (2021). *Co-production and educational change: Moving from formulaic to flexible scaling strategies* (pp. 270–289). https://doi.org/10.4018/978-1-7998-4975-9.ch015.

Karlsen, L., & Lekhal, R. (2019). Practitioner involvement and support in children's learning during free play in two Norwegian kindergartens. *Journal of Early Childhood Research, 17*(3), 233–246. https://doi.org/10.1177/1476718x19856390.

Kellogg Foundation, W. K. (2004). *Using logic models to bring together planning, evaluation, and action: Logic model development guide*.

Lenes, R., McClelland, M. M., ten Braak, D., Idsøe, T., & Størksen, I. (2020). Direct and indirect pathways from children's early self-regulation to academic achievement in fifth grade in Norway. *Early Childhood Research Quarterly, 53*, 612–624. https://doi.org/10.1016/j.ecresq.2020.07.005.

McClelland, M. M., Acock, A. C., & Morrison, F. J. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school, 2006/10/01/ *Early Childhood Research Quarterly, 21*(4), 471–490. https://doi.org/10.1016/j.ecresq.2006.09.003.

McClelland, M. M., Cameron, C. E., Duncan, R., Bowles, R. P., Acock, A. C., Miao, A., et al. (2014). Predictors of early growth in academic achievement: The head-toes-knees-shoulders task [Original Research], 2014-June-17 *Frontiers in Psychology, 5*(599). https://doi.org/10.3389/fpsyg.2014.00599.

McClelland, M. M., Cameron, C. E., McDonald Connor, C., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology, 43*(4), 947–959.

Ministry of Education and Research. (2017). *Framework plan for the content and tasks of kindergartens*. https://www.udir.no/globalassets/filer/barnehage/rammeplan/framework-plan-for-kindergartens2-2017.pdf.

Moser, T., & Martinsen, M. T. (2010). The outdoor environment in Norwegian kindergartens as pedagogical space for toddlers' play, learning and development, 2010/12/01 *European Early Childhood Education Research Journal, 18*(4), 457–471. https://doi.org/10.1080/1350293X.2010.525931.

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, Rimes, Vocabulary, and Grammatical Skills as Foundations of Early Reading Development: Evidence From a Longitudinal Study. *Developmental Psychology, 405*, 665–681. https://doi.org/10.1037/0012-1649.40.5.665.

Osborne, S. P., & Strokosch, K. (2013). It takes two to tango? Understanding the Co-production of public services by integrating the services management and public administration perspectives. *British Journal of Management, 24*(S1), S31–S47. https://doi.org/10.1111/1467-8551.12010.

Pianta, R. C. (1999). *Enhancing relationships between children and teachers*. American Psychological Association.

Pontoppidan, M., Keilow, M., Dietrichson, J., Solheim, O. J., Opheim, V., Gustafson, S., et al. (2018). Randomised controlled trials in Scandinavian educational research, 2018/07/03 *Educational Research, 60*(3), 311–335. https://doi.org/10.1080/00131881.2018.1493351.

Rege, M., Solli, I. F., Størksen, I., & Votruba, M. (2018). Variation in center quality in a universal publicly subsidized and regulated childcare system. *Labour Economics, 55*, 230–240. https://doi.org/10.1016/j.labeco.2018.10.003.

Rege, M., Størksen, I., Solli, I. C. F., Kalil, A., McClelland, M. M., Ten Braak, D., … Hundeland, P. S. (2019). Promoting Child Development in a Universal Preschool System: A Field Experiment. *Cesifo Working Papers*, 7775.

Sabol, T. J., Soliday Hong, S. L., Pianta, R. C., & Burchinal, M. R. (2013). Can rating Pre-K programs predict children's learning?. August 23, 2013 *Science, 341*(6148), 845–846. https://doi.org/10.1126/science.1233517.

Schindler, H. S., Fisher, P. A., & Shonkoff, J. P. (2017). From innovation to impact at scale: Lessons learned from a cluster of research–Community partnerships, 2017/09/01 *Child Development, 88*(5), 1435–1446. https://doi.org/10.1111/cdev.12904.

Schleicher, A. (2019). *Helping our youngest to learn and grow*. https://www.oecd-ilibrary.org/content/publication/9789264313873-en.

Størksen, I., & Mosvold, R. (2013). *Assessing early math skills with tablet computers: Development of the Ani Banani Math Test (ABMT) for young children Utdanning 2020.* Oslo: The Norwegian Research Council.

Størksen, I., Ellingsen, I. T., Tvedt, M. S., & Idsøe, E. M. C. (2013). Norwegian vocabulary test (NVT) for children in transition from kindergarten to school: a psychometric evaluation of a computer tablet test. [Norsk vokabulartest (NVT) for barn i overgangen mellom barnehage og skole: Psykometrisk vurdering av en nettbrettbasert test.]. *Spesialpedagogikks forskningsdel, 04/13*, 40–54.

Størksen, I., Braak, tenD., Breive, S., Lenes, R., Lunde, S., Carlsen, M., … Rege, M. (2018). Playful learning - a science based preschool program from the Agder project. [Lekbasert læring - et forskningsbasert førskoleopplegg fra Agderprosjektet.]. *GAN Aschehoug*.

ten Braak, D., & Størksen, I. (2021). Psychometric properties of the Ani Banani Math Test. *European Journal of Developmental Psychology*, 1–18. https://doi.org/10.1080/17405629.2021.1879046.

ten Braak, D., Størksen, I., Idsoe, T., & McClelland, M. (2019). Bidirectionality in self-regulation and academic skills in play-based early childhood education. *Journal of Applied Developmental Psychology, 65*, Article 101064. https://doi.org/10.1016/j.appdev.2019.101064.

Thornton, P. K., Schuetz, T., Förch, W., Cramer, L., Abreu, D., Vermeulen, S., et al. (2017). Responding to global change: A theory of change approach to making agricultural research for development outcome-based, 2017/03/01/ *Agricultural Systems, 152*, 145–153. https://doi.org/10.1016/j.agsy.2017.01.005.

Tymms, P., Merrell, C., Thurston, A., Andor, J., Topping, K., & Miller, D. (2011). Improving attainment across a whole district: School reform through peer tutoring in a randomized controlled trial, 2011/09/01 *School Effectiveness and School Improvement, 22*(3), 265–289. https://doi.org/10.1080/09243453.2011.589859.

Wechsler, D. (2004). *The Wechsler intelligence scale for children* (fourth ed.). Pearson.

Wegener, D. T., & Blankenship, K. L. (2019). Ecological validity. *Encyclopedia of social psychology social psychology (general)*. http://sk.sagepub.com/reference/socialpsychology/n167.xml.

Weisberg, D. S., Hirsh-Pasek, K., & Golinkoff, R. M. (2013). Guided play: Where curricular goals meet a playful pedagogy. *Mind Brain and Education, 7*(2), 104–112. https://doi.org/10.1111/mbe.12015.

Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., Gormley, W. T., et al. (2013). *Investing in our future: The evidence base on preschool education*. https://www.fcd-us.org/assets/2016/04/Evidence-Base-on-Preschool-Education-FINAL.pdf.