On the Trail of Early Numeracy Skills

Understanding, identifying and ameliorating young children's early numeracy skills: A multimethod approach

Anita Lopez-Pedersen

Department of Special Needs Education

Faculty of Educational Sciences

University of Oslo

2020

© Anita Lopez-Pedersen, 2020

Series of dissertations submitted to the
Faculty of Educational Sciences, University of Oslo
No. 321

ISSN 1501-8962

All rights reserved. No part of this publication may be
reproduced or transmitted, in any form or by any means, without permission.

Cover: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.

Abstract

The development of well-functioning numeracy skills is a foundation for further mathematical skills. Mathematical skills develop in a cumulative fashion, and to help children establish these skills, we need to better understand how they in fact develop, how to identify children who are at risk of developing mathematical learning difficulties, and last but not least – how we can help remedy these difficulties. To better understand and ultimately support these children, we need to apply different methods and research designs. We need to make sure we raise questions that are detailed enough in order to avoid brevity, and help to answer the major question which is – which are the most important skills to help children who struggle in mathematics? Which particular skills might be more relevant to be assessed in order to identify children who are at risk of developing mathematical learning difficulties?

The first study examined the developmental relationship between the approximate number system and early mathematical skills, in two different datasets. The main objective was to further investigate the theory that the approximate number system has a potential casual influence on mathematical development. First, we reanalyzed the dataset from a recent study by Elliott, Feigenson, Halberda, and Libertus (2019). Using cross-lagged panel model Elliott et al. (2019) claimed a reciprocal relationship between the approximate number system and early mathematics, however when reanalyzing this dataset by using a novel methodological approach, a random-intercept cross-lagged panel model, no evidence of a reciprocal relationship was found. Second, in a 1-year longitudinal study with three time points we examined the developmental relationship of the approximate number system and addition skills with the same methodological approach as in the reanalysis of Elliott et al.'s (2019) data. Here, the results did not show any evidence supporting a reciprocal relationship between the approximate number system and mathematical development either. Combined, this questions the idea that the approximate number system plays a vital role in the development of early mathematics, and vice versa. Moreover, this study displayed how different methodological approaches lead to different results.

The second study is a validation study. The psychometric properties of the Early Numeracy Screener that was developed from a theoretical model, the core numerical skills model by Aunio and Räsänen (2016). The Early Numeracy Screener aims to detect children who are at risk of developing mathematical learning difficulties later on, and furthermore aims to identify three sets of early numeracy skills; namely counting skills, numerical relational skills, and basic arithmetic skills. Confirmatory factor analysis found evidence for a

three-factor model, establishing construct validity. Furthermore, criterion-related validity was found in crosstabulation and correlation with a national test taken towards the end of the school year. The results indicated that the Early Numeracy Screener identifies three dimensions of early numeracy skills, and hence serves as functional screener for first graders in mathematics.

The third study is a randomized controlled trial of low-performing first graders. This study aimed to improve early numeracy skills and evaluated the effects of a 14-week early numeracy program designed to boost numeracy skills of low-performing children at risk of developing difficulties in mathematics. The intervention targeted counting skills, numerical relational skills, and basic understanding of arithmetic. The intervention produced modest benefits ($d = 0.20$) on early numeracy (counting and numerical relational skills) learning, but those were not significant. There were moderate and reliable effects on word problem solving ($d = 0.41$); however, the effects were reduced and faded at the second follow-up test (after a second intervention phase of 6 weeks with intervention training once per week) and at follow-up test (6 months after the intervention) compared to the immediate post-test.

Combined, the findings from these three studies have attempted to ultimately enhance the understanding of what influences mathematical development in young children, what early mathematical skills generate mathematical development. More precisely, these studies indicated that the debated approximate number system did not have a strong as effect on mathematical development as previous studies suggest. Additionally, in order to make fine-grained assessment when children are in the early development of mathematics these studies found that early numeracy skills could be divided into three separate numeracy skills. Knowledge about what influences mathematical development and the notion that subskills in early numeracy could be identified, furthered skills that could be targeted in interventions supporting struggling learners in mathematics. All three studies have implications for future research and for the practice field. For future research, this thesis have highlighted how novel and innovative methodological approaches can affect the evidence in which inferences are drawn. Furthermore, this thesis have implications for the Norwegian practice field and provide Norwegian schools and teachers an assessment tool to help identify children at risk of developing learning difficulties in mathematics, and with intervention material to help support these children.

Acknowledgements

Primarily I want to thank all the participants - teachers, children, and parents. An enormous thank you to the schools for trusting me with your teachers' time set aside to conduct the screener and intervention. Having worked as a teacher for many years myself, I know how putting that much time and effort into this cost. Teachers – thank you for you the hard and dutiful work you put into this. A special thanks to all the participating children. It has been a joy meeting some of you throughout this process; it has been motivating and an important reminder of whom this thesis really is about.

Thank you to my supervisors, I could not have wished for a better team. Monica Melby-Lervåg, it has been such an honor being under your wings. I have learned so much, both methodologically, substantially, but also how to be a better scientific writer. You truly are an inspiration. Riikka Mononen, I am so thankful for all your knowledge and wisdom. You are both strict and friendly, a combination I have been incredibly fond of. Thank you for the cool side projects we have had together outside the work with this thesis - kiitos. Pirjo Aunio, I am so lucky to harvest the fruits of your experience as a supervisor in this field, thank you for sharing your knowledge and challenging me with asking tough (but necessary and eye opening) questions.

I want to thank the Department of Special Needs Education for giving me the opportunity to do this thesis. A special thanks to former head of the department, Berit Rognhaug, for making a game changing phone call to me the spring of 2015. I also wish to thank the two research groups I have been in over the course of these years, CLL and LinCon. It has been a great learning experience working closely with admirable researchers with such a range in research fields and methodology.

A special thanks to my additional co-authors, Arne Lervåg, Ronny Scherer, and Johan Korhonen. Thank you for exposing me to and sharing your immense methodological knowledge, I have the utmost respect for you expertize. I consider you SEM-magicians, SEM-heroes, and fortunately you were all great SEM-mentors.

Furthermore, I want to direct a huge thanks to Silke Göbel for useful, motivating, and most appreciated comments on the final review.

Thank you also to Østfold University College for trusting me with offering me a job long before I finished my PhD. A special thanks to Rune Andreassen for your considerate and invaluable support in the finishing line of this thesis.

Huge thanks to colleagues at the department. Thank you Katzy and Marika for great support in the research administration. Especially the supreme PhD-students for such a great working environment. Thank you for making every day at work fun, and thank you for constant encouragement, laughter and making celebratory events, no excuse too small. A special mention to Siri, Stine, Hanne, Anne, Arne, Linn, Anette, Gøril, Silje Hok., Terje, Veerle and Jannicke. Extra thanks to my partner in crime, Mia Heller – it has been an amusing journey sharing these four years along your side, from finish to start. In addition, thank you to my office neighbor Christiane L. Haukedal for our reciprocal relationship (pun intended) when it comes to sharing podcast tips. Unwinding with interesting and entertaining listening after long working days should not be underestimated, especially not in the last inning of the PhD.

Thank you to friends and family. Huge thanks to Marie, Kristine, and Liv H. for exhilarating non-work related escapades. Thank you mum for getting up five in the morning allowing me to go to work ridiculously early. Thank you dad for always asking me «hvordan går det med den forskningen, er du i rute? », both showing an endless interest but also reminding me not to waste any time. Thank you both for taking such good care of my girls throughout these years, you are indeed extraordinary grandparents. Thank you to my sister Berit and her family – thank you Berit for taking me on extended weekends to various Balkan cities throughout these course of these years, batteries have indeed been recharged. Huge thanks to Kjell. I would not have been here if it had not been for you. You were the first one to motivate and encourage me to enter academia, and your unconditional confidence in me means more than word can express. Thank you for stimulating conversations, always reminding me why research in this field is imperative.

Finally, the greatest girls in my world. Thank you Signe for every day asking me how my workday has been and taking an interest in what I do (even though gazing eyes reveal that you might find my answers boring and abstract at times), and for holding the fort being the elder sister, your care for Liv and me is incomparable. Thank you Liv for giving the best hugs and writing small letters always lifting my spirits. Signe and Liv – I am so grateful for all time-spent hiking in the forest, slalom skiing, playing football during these years and thank you for never getting tired of watching Harry Potter with me. To paraphrase Bess Streeter Aldrich, there is neither subtraction nor division in my love for you, only addition and multiplication.

PART 1

Extended abstract

Contents

1 Introduction

Efforts to support children's numeracy skills are of long-term importance to children's later academic success. We need to understand the origins and complexity of early numeracy skills in order to identify children who are at risk for developing difficulties with mathematics and remedying these difficulties as early as possible. The main topic of this thesis is early numeracy, subdivided into the development of early numeracy skills, assessment of early numeracy skills, and educational support targeting early numeracy skills. Three studies are presented:

The approximate number system and early mathematics: the developmental relationship. The first study presents an analysis of two longitudinal datasets about the developmental relationship between approximate number system and early mathematics.

Validation of an early numeracy screening tool. The second study is a validation of a screening tool, the Early Numeracy Screener for first graders, which can be used as a means by which to detect children struggling with early numeracy and who are therefore at risk of developing learning difficulties in mathematics.

A randomized controlled trial. The third study presents a randomized controlled trial with the aim of ameliorating low-performing first graders' early numeracy skills.

Pursuing the objective of understanding and supporting mathematical thinking and learning relies on different conceptions of fundamental theoretical constructs (Alock, Gilmore, & Inglis, 2013), as well as methodological approaches. There are many unanswered questions in this field, in terms of theories about how we create and develop mathematics and how components of this development are associated with each other – if indeed, they are associated at all. The initial question at the outset of this thesis was (an enormous) "How?" – How can we support children who struggle in learning mathematics? Several other queries are necessary to respond to this "how", for instance "who" - Who needs this support, which children? Why do they need exactly this support and in what manner? In addition, on what grounds do we decide how to provide support, why do we prioritize some components of mathematical development over others in both identifying children at risk for mathematical difficulties and providing support. Are some components in mathematical development more important than others? The enormous question "how" quickly evokes more questions, and the questions become more and more concrete and detailed, while at the same time they are entangled. Answering these seemingly easy questions raises even more questions. We need more knowledge when it comes to finding key numeracy skills for mathematical development as this insight might have consequences for developing valid and reliable measures and designing effective interventions.

2 The path to mathematical competence

It has been well established that children's quantitative competencies when they enter formal schooling predict their relative mathematics achievement throughout schooling (Duncan et al., 2007; Ritchie & Bates, 2013). Their mathematical competencies when they leave school predict their employability and wages throughout adulthood (Brynner, 1997; Rivera-Batiz, 1992). This theoretical chapter opens with an overarching theoretical model on how we can understand learning difficulties in general. The objective for including this all-embracing model is to place mathematical development in a broader perspective and hence lead the way to theoretical models and frameworks on mathematical development. The models depicted are frameworks on how we both can understand typical development and learning difficulties in mathematics. The models discussed in this chapter are already existing models, and are challenged methodologically in terms of inferences drawn from these "box-and-arrows" models. These theoretical models on mathematical development will submit to some of the unresolved issues related to causality in the field of mathematics. Furthermore, in this chapter, the different numeracy components will be presented. Additionally, this chapter on theory includes matters such as mathematical learning difficulties, assessment of early numeracy, and early numeracy interventions.

2.1 Attempts to explain mathematical development

It has been suggested that quantitative abilities provide the foundation for the emergence of more complex numeracy, counting, and arithmetic skills during the preschool years (Gelman & Gallistel, 1978). Generally, early numeracy is typically defined as the understanding of numbers prior to formal instruction (Howell & Kemp, 2010), for instance learning number-words, its sequence and the acquisition of counting skills connecting the number word to quantity. Studies have related children's mathematical achievement to specific aspects (e.g. counting skills) of their early numerical competencies (Hannula-Sormunen, Lehtinen, & Räsänen, 2015). Developing well-functioning early numeracy skills is a groundwork for later mathematical skills. It is considered a vital pathway to employment in a society in which the demands for mathematical reasoning, problem solving and being able to analyze information is steadily increasing (Geary, Bailey, & Hoard, 2009; Trilling, & Fadel, 2009). In this thesis, the concepts, numeracy and mathematical skills and abilities, are used. The umbrella term *mathematical development* and *mathematical skills* encompasses numeracy skills, and numeracy skills, as we will see, are thought to be a building block in theories of mathematical development.

A cognitive explanation of learning difficulties focuses how skills are learned and how typical development might be disturbed (Hulme & Snowling, 2009). A complete explanation of any learning difficulty will involve several layers of description. Morton and Frith (1995) argued that it would be useful to make unambiguous diagrams for theoretical explanations. Figure 1 is an example of such a general model and depicts several sections of learning. Therefore, it is hypothesized that there are separable modules that underlie particular skills, and these subsystems have different functions. To exemplify, the model might help describe the association between the cognitive level (e.g. working memory) that is hypothesized to affect the behavioral level in mathematics (e.g. arithmetic fact retrieval). The model, however, does not assume an opposite causal link; the debate concerning the causal link in mathematics is elaborated further later in this thesis.

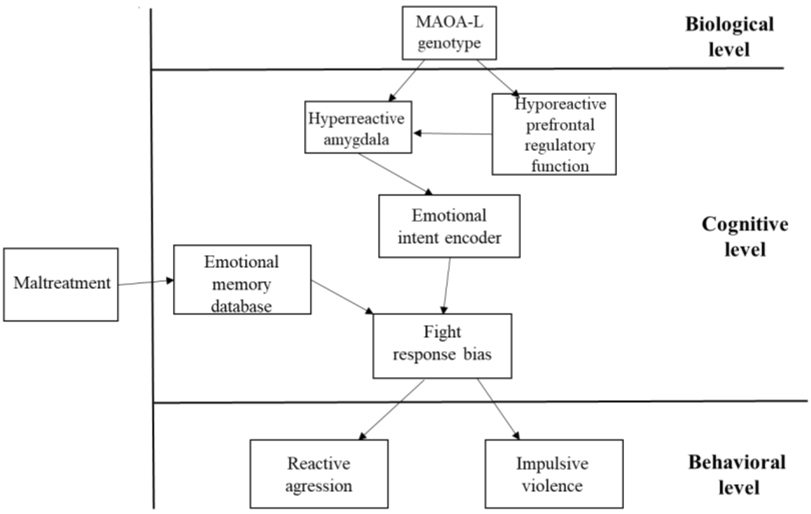

Figure 1. A model of the potential gene-brain-cognition-behavior pathways from genotype MAOA-L to reactive aggression. Adapted from Morton and Frith (1995)

Although some aspects of development are subject to genetic influences, it is important not to adopt a deterministic mindset; genetic influence does not undermine the importance of environment; hence, providing support for struggling learners is imperative. It has been demonstrated that for mathematics, heritability (inherited genetic factors) explains a large variation in children's skills, around 60-70% (de Zeeuw, de Geus & Boomsma, 2015). This does not imply that mathematic skills are unalterable since 30-40 % consequently would then not be under genetic control. Furthermore, it also seems to vary by country, for instance De Zeeuw et al., (2015), compared with Hart, Petrill, Thompson, and Plomin (2009) displaying that genetic factors explained more of the variance in the U.S. than in the UK. This does however imply that it can be difficult to increase the mean and reduce the variation in a group at the same time. Similarly important is the neurological perspective on the development of learning and learning difficulties. In the field of mathematics, neurological research is essential to understanding development through investigation of brain activation patterns and, for example, which brain areas involved in typical development and functioning might be disturbed both in selected and unselected samples. In-depth discussions of the genetic and neurological domains of mathematical learning difficulties are however not included in this thesis. Although these levels are a foundation and of course important, this thesis operates on a cognitive and behavioral level because to understand how to remedy mathematical difficulties, it is necessary to understand the development and interaction between the cognitive and the behavioral level. This does not diminish the importance of neuroscientific and genetic studies, but simply because this was neither the aim nor the rationale of this thesis.

2.2 Theories concerning factors that influence numeracy development

Which aspects of thinking originate from instruction, and which emerge independently of specific experience? We need to more fully understand the factors that scaffold early mathematics development. To do this, we need to look more closely at some of the models of mathematical development. Researchers have used information from a variety of sources to address the question of how children acquire skills in conventional mathematics, including longitudinal studies of typically developing children (e.g. Aunola, Leskinen, & Nurmi, 2004; Bull, Espy, & Wiebe, 2008; De Smedt, Verschaffel, & Ghesquière, 2009; Jordan, Kaplan, Locuniak, & Ramineni, 2007) and research with children who have severe mathematical learning difficulties (i.e. developmental dyscalculia, e.g. Berch & Mazzocco, 2007). Most of the studies in this field, however, are cross-sectional, which is a critical issue when it comes to studying development. Therefore, in this thesis, longitudinal studies will be prioritized since they can be used to generate causal hypotheses.

2.2.1 The triple code model of numerical processing

In 1992, Dehaene proposed a model, namely the triple code model and this model has become a widespread model in the field of number processing. Dehaene's triple code model linked the external, encoded representation with different internal mental codes, so that information represented as digits activated representations that differed from those activated by words or quantities. An unambiguous distinction has emerged between a system of calculation procedures based on the quantity on the one hand and a second calculation system relying on memorized facts (Dehaene, 1992; Dehaene & Cohen, 1991). Incorporating this idea, see Figure 2, (Dehaene 1992; Dehaene & Cohen, 1995; Dehaene & Cohen, 1997), the triple code model is a multirouted model of numerical processing postulating three functionally sovereign but related codes (Dehaene, 1992).

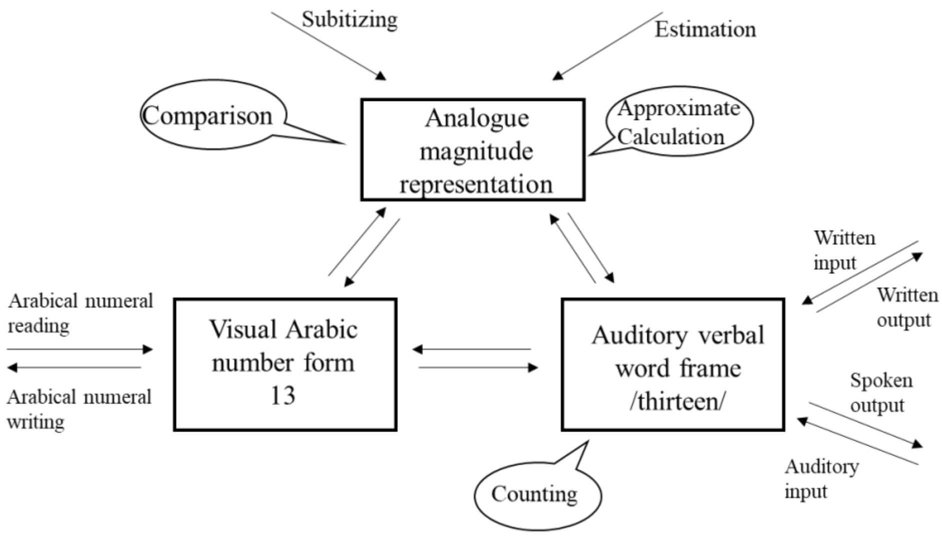

Figure 2. Triple Code Model of Numerical Cognition with elements related to early numeracy, adapted from Dehaene (1992)

The triple code model can be seen as an operationalization of early numeracy, and has a detailed numerical aspect of early numeracy. The model is theorized by three premises that must occur. Numbers may be represented mentally in three main representations of numbers (Dehaene & Cohen, 1997): a visual Arabic code, an analogical magnitude code, and a verbal code. The analogue magnitude representation is knowledge about the proximity and relative size of quantity. The foundation of the analogue magnitude representation is a preverbal system, suggesting that children have an innate understanding of magnitudes (Dehaene, 1992) that can be measured in preverbal infancy. Linguistic numerical tasks, such as retrieving answers to arithmetic combinations, activate a region that has been linked to language processing, including reading and phoneme detection. The triple code model sets these abilities into three groups according to the format in which numbers are manipulated (Dehaene & Cohen, 1997). The bidirectional arrows between the three codes suggests the three codes being both independent, but at the same time interdependent and linked together.

The original version of the triple code model and subsequent extensions primarily relied on theoretical considerations derived from arithmetic deficits following neuropsychological impairments such as an acquired deficit in calculation (Dehaene & Cohen, 1997). These models were derived from investigations involving adults; we therefore need to take a closer look at studies involving younger children. At present, several early predictors for children's early mathematical attainment are predominant in the literature; visuospatial skills (Butterworth 2005); mental number line precision (Praet & Desoete, 2014; Reeve, Paul, & Butterworth, 2015); innate number competence (Duncan et al., 2007; Jordan, Kaplan, Ramineni, & Locuniak, 2009); counting skills (Jordan et al., 2007; Muldoon, Towse, Simms, Perra, & Menzies, 2013); and linguistic skills (Praet, Titeca, Ceulemans, & Desoete, 2013). In 2018, Siemann and Petermann assessed the literature containing contradictory findings on how the triple code model can be applied to mathematical development, and challenged what they considered the model's unanswered questions. First, Siemann and Petermann (2018) stated that developmental trajectories are not clearly expressed for the model. Second, it was suggested that the many impacts of domain-general factors are not accounted for.

Support for the triple code model has however been found in recent studies (Malone, Burgoyne, & Hulme, 2019; Malone, Heron-Delaney, Burgoyne, & Hulme, 2019). A latent variable path model showed that early number knowledge and numerosity discrimination were unique predictors of arithmetical development (Malone, Burgoyne, et al., 2019), and another study showed that tasks involving mapping magnitudes onto verbal or visual stimuli predicted arithmetic performance over and beyond predictors such as age and IQ (Malone, Heron-Delaney, et al., 2019).

2.2.2 Pathways to Mathematics

In 2010, LeFevre et al. (2010) proposed and tested a slightly different model compared with Dehaene's (1992) model of mathematical development in children's acquisition of mathematics. This pathway to mathematical development model hypothesized by LeFevre and colleagues (2010) consists of three independent cognitive precursors or pathways: linguistic, quantitative, and spatial attention. The pathways model posits that linguistic skills should predict children's performance in early numeracy, such as naming numbers or writing Arabic digits. Language as a measure has recently been stressed in the prediction of numeracy development (Purpura, Hume, Sims, & Lonigan, 2011; Sarnecka, Kamenskaya, Yamana, Ogura, & Yudovina, 2007). Studies do reveal contradictory findings, however; some studies (Barner, Chow, & Young, 2009; Negen & Sarnekca, 2012) found that general measures of language development predict number and number word knowledge, although other studies did not support such a link (Ansari et al., 2003). Praet and colleagues (2013) found that expressive language explains 24 % of the variance in arithmetic skills, even after controlling for number naming and procedural counting. This is in line with other studies (Purpura et al., 2011), and in accordance with studies finding similar patterns of growth in linguistic and quantitative skills (Jordan, Kaplan, Olàh, & Locuniak, 2006; Jordan et al., 2007). Different language components might contribute to different mathematical skills (e.g. oral language, receptive language, understanding of grammatical rules, Storch & Whitehurst, 2002). Still other studies have found language skills to be linked with children's word problem solving skills (Fuchs et al., 2008; Fuchs et al., 2010). Linguistic skills, however, did not predict nonlinguistic arithmetic (LeFevre et al., 2010). Nonlinguistic arithmetic were arithmetic tasks requiring children to represent and mentally manipulate quantities, but do not require labelling those quantities with numbers or link to Arabic symbols (LeFevre et al., 2010). This supports the proposed distinction between these pathways, leaving numerically relevant linguistic and quantitative skills as two distinct factors. The model depicted in Figure 2 displays the predictions premising the model.

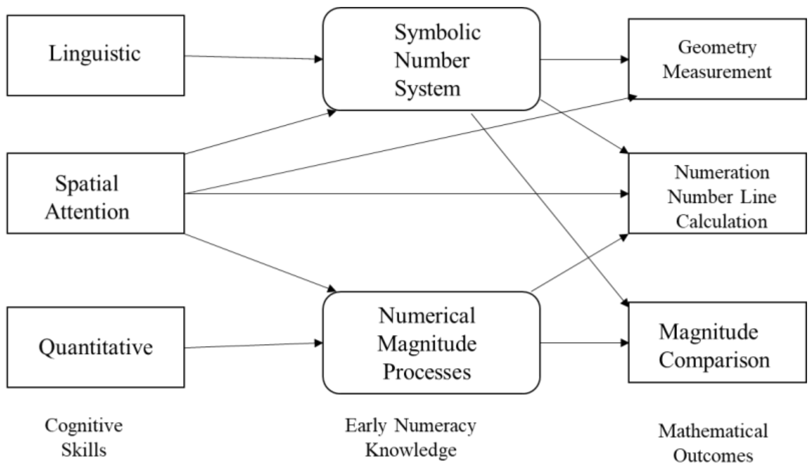

Figure 2 Pathways to Mathematics, adapted from LeFevre et al. (2010)

LeFevre et al. (2010) suggested that future work is needed to refine and develop the pathways model theoretically, empirically, and methodologically. The pathway model did not address the question of how knowledge from linguistic and quantitative pathways is integrated; although it was suggested that working memory might play an important role in this process (Krajewski & Schneider, 2009; Swanson, Jerman, & Zheng, 2008).

In 2015, Sowinski et al. attempted to refine the pathways to mathematics model. They tried to improve the quantitative pathway by combining children's subitizing, counting, and symbolic magnitude comparison skills, and furthermore hypothesized that quantitative, linguistic, and working memory pathways could account for unique variance in the numerical outcomes. However, the results indicated that only the quantitative and linguistic pathways, not working memory, accounted for unique variance. Although the refinement of this model supported the 2010 model empirically, the study had the limitation of being cross-sectional, not longitudinal, as was the original proposal of the pathways to mathematics model (LeFevre et al., 2010).

In 2017, Lira, Carver, Douglas, and LeFevre attempted yet again to integrate mapping among non-symbolic quantities, i.e. between spoken number words and written digits into the model. Empirical support for this integration was found in previous studies displaying the use of written symbols predicting arithmetic skills (Zhang et al., 2014). Furthermore, Lira et al. (2017) investigated to find a way this unit might be mediated by counting sequence knowledge suggested by (Purpura, Baroody, & Lonigan, 2013) and other numerical language skills (Moll, Snowling, Göbel, & Hulme, 2015). After using latent variables testing their hypothesis, Lira et al. (2017) found that children's verbal counting predicted their knowledge of the number symbols and their ability to understand and manipulate exact quantities. In the integrated model, the various precursor and mapping tasks were arranged in a linguistic/symbolic and quantitative pathway consistent with the original model from 2010. The study concluded that children's knowledge about non-symbolic exact quantities, spoken number word, and digits predicted their ability to map between symbolic and non-symbolic exact quantities. The mappings between written digits and non-symbolic exact quantities developed later than the other mappings. Thus, Lira et al. (2017) claimed this to be evidence for a model of early number knowledge in which integration across symbolic and non-symbolic representations of exact quantity underlies the development of children's number comparison skills.

2.2.3 Core numerical skills model

Another theoretical model that has been suggested to explain the foundation of early numeracy skills is the core numerical skills model. Aunio and Räsänen (2016) hypothesized a model of crucial numerical factors for the development of mathematical skills among children aged five to eight years old. Their model was based on results from longitudinal studies, more precisely mathematical subskills that predict later mathematics (e.g. Aunola et al., 2004; LeFevre et al., 2010). In order to find further support for these construct, they qualitatively analyzed normed test batteries intended to measure the development of mathematical skills, the idea being to investigate how these groups of skills were operationalized within different test batteries. The core numerical skills model (see Figure 3) divided skills into four parts, symbolic/non-symbolic number sense, understanding mathematical relations, counting skills, and basic skills in arithmetic (Aunio & Räsänen, 2016).

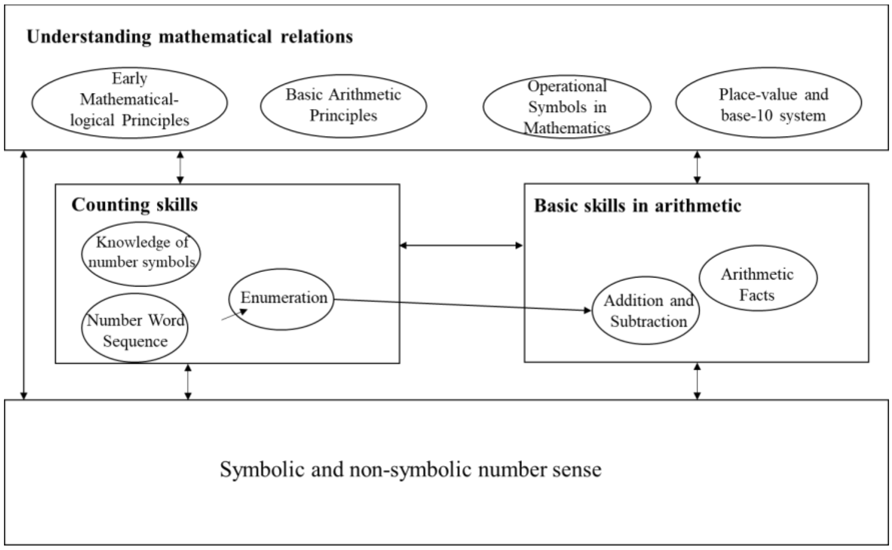

Figure 3 Core numerical skills model, adapted from Aunio and Räsänen (2016)

In the core numerical skills model, symbolic and non-symbolic skills are defined as processes where approximate evaluation of magnitudes or symbols representing magnitudes are used. Understanding mathematical relations in the core numerical skills model include domains that constitute early numeracy skills; namely, understanding numerical relations (e.g. understanding of the quantitative and non-quantitative relationships between the elements in the task). Furthermore, numerical relational skills include mathematical-logical principles, understanding the meaning of the base-10 system (Aunio & Räsänen, 2016). A third component of the core numerical skills model is counting skills. Counting skills refers to the child's knowledge of number symbols, skills in moving within the sequence of the number words and enumeration (Aunio & Räsänen, 2016). Basic arithmetic skills constitute the fourth component, and in 5–8-year-olds, this refers to the degree to which a child masters mainly addition and subtraction tasks with number symbols (Aunio & Räsänen, 2016). Visually, the model's boxes and arrows display that non-symbolic and symbolic number sense is the foundation the components counting skills, numerical relational skills, and basic arithmetic skills are based on, it but also implies this association to be bidirectional.

Although the core numerical skills model was designed without being empirically tested prior to being published, one can find empirical support for the four components' predictive influence of early numeracy and mathematical development. Longitudinal studies have established that numerical relational skills are a central part of early numeracy development (Aunio & Niemivirta, 2010; Desoete, Stock, Schepens, Baeyens, & Rieyers, 2009; Stock, Desoete, & Roeyers, 2009). Counting strategies are, perhaps not surprisingly, an imperative aspect of children's early numerical knowledge (Wright, Martland, & Stafford, 2006). Notably, the skills displayed in the model are supported by previous research describing predictors of early numeracy (Desoete, Ceulemans, De Weerdt, & Pieters, 2012; Gersten et al. 2012; Moeller, Pixner, Zuber, Kaufmann, & Nuerk, 2011). Additionally, this was not the first nor last attempt to create multifactor models for early numeracy and mathematical development (e.g. Cirino, 2011; Hirsch, Lambert, Coppens, & Moeller, 2018).

In 2019 Aunio and colleagues empirically tested the core numerical skills model using a cross-sectional design with confirmatory factor analysis, before it was tested yet again as part of a validation for the Early Numeracy Screener (see Study 2). The factors for counting skills, numerical relational skills, and basic arithmetic skills were identified as three separate dimensions (Aunio et al., 2019). However, the model has not yet been tested in a longitudinal design and this raises a question of dimensionality. In Study 2, three of the four components in the core numerical skills model were empirically tested, trying to establish whether the model indeed consisted of several factors or was unidimensional. Even though the analysis favored the three-factor model, the model fit for a one-factor model was also good, and arguably indicated the model might be unidimensional. Thus, even if the factors that create the foundation for mathematic skills are conceptualized as a set of different skills, these skills might be highly related. For further discussion, see Study 2.

2.2.4 Conceptual and procedural understanding

Another way of understanding the foundation of mathematical skills is suggested to be based on a development of both a conceptual and a procedural understanding of the task. Conceptual knowledge refers to the implicit or explicit understanding and is a flexible knowledge that is not tied to specific types of task. Procedural understanding is the ability to execute action sequences to solve problems (Rittle-Johnson, Siegler, & Alibali, 2001). According to this theory, to develop well-functioning mathematical skills, a child has to master both conceptual and procedural understanding.

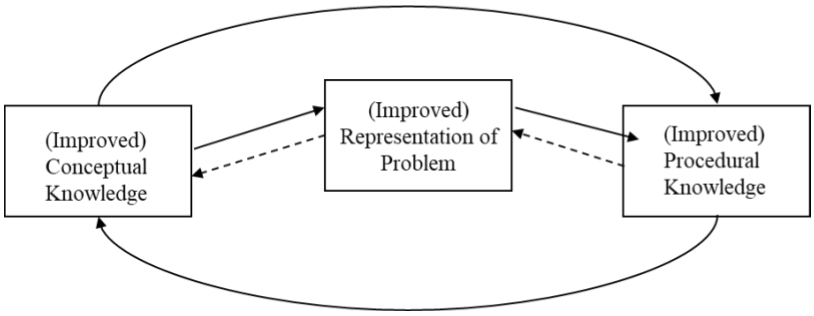

Figure 4. Iterative model for the development of conceptual and procedural knowledge, adapted from Rittle-Johnson et al. (2001).

As for empirical support, it has been demonstrated that children's conceptual and procedural knowledge contribute to their competencies in an area of mathematics, such as arithmetic (Rittle-Johnson & Siegler, 1998; Rittle-Johnson et al., 2001). These two types of knowledge (conceptual and procedural) lie on a continuum and cannot always be separated (Rittle-Johnson et al., 2001). According to this theory, depicted in Figure 4, children's conceptual knowledge of the magnitude of numbers is especially important because it involves the ability to estimate the magnitude of numerals and sets of objects (Siegler, 2016; Siegler & Booth, 2005). Fundamental arithmetic includes both procedural knowledge (e.g. 7 – 5 = 2, knowing how to execute the task) and conceptual knowledge (e.g. 7 – 5 = 2, implicit and explicit understanding, knowing that subtraction is inverse addition, and one can use addition to solve this task (Bisanz & LeFevre, 1990).

What weakens the theory of procedural and conceptual knowledge is that they seem to develop hand in hand, and are highly interweaved. This was also addressed by Rittle-Johnson, et al. (2001) themselves, they highlighted that the bidirectional relations between procedural and conceptual knowledge might lead to the iterative development of the two types of knowledge. Thus, it seems rare that a child is able to develop procedural knowledge without having conceptual knowledge or vice versa. Notably, the research on conceptual and procedural knowledge led to the development of an important line of research that focused on strategies of arithmetic facts and problem solving. A strategy is a procedure that is involved in a flexible, goal-oriented manner and that influences the selection and implementation of subsequent procedures (Bisanz & LeFevre, 1990). For example, there are individual differences in memorizing arithmetic facts. Children with mathematical learning difficulties struggle more when retrieving arithmetic facts from memory than children with typical development (Geary, 2010; Jordan, Glutting, & Ramineni, 2008). Children who perform poorly in arithmetic rarely use retrieval of facts and tend to rely on back-up strategies (e.g. counting fingers instead of retrieving the answer directly) for solving even basic arithmetic tasks (Geary, 1993; Ostad, 1997,1998; Siegler, 1988). Children without mathematical difficulties show increased reliance on retrieval strategies and a decreased reliance on back-up strategies (Ostad, 2000; Ostad & Sorensen, 2007). In a sense, being rigid in strategy use and having only few strategies for problem solving can be seen as an adjoining cause of mathematical difficulties. However, it could also be a consequence of the difficulties. It can perhaps be best understood as a characteristic rather than a cause or consequence.

2.3 Domain-general and domain-specific factors

Mathematical proficiency is determined by many cognitive factors (e.g.Geary, 1993), and a common way to distinguish between them is to distinguish between domain-general (e.g. working memory) versus domain-specific (e.g. counting) skills. It has been argued that the theoretical models discussed in the previous section can be labelled domain-specific models since they explain mathematical development mainly as a result of development in skills that are close to mathematics (e.g. Aunio & Räsänen, 2016).

The current understanding of the terms domain-general and domain-specific factors has been shaped by discussions about whether there are domain-specific modules in the mind (Fodor, 1983) and whether infants enter the world with innately pre-specified core knowledge (Dehaene, 2001). Fodor (1983) defined the modules need to be innately pre-specified, which is line with the hypothesis that infants possess innately pre-specified domain-specific core knowledge (Dehaene, 2001). For example, Feigenson, Dehaene, & Spelke (2004) postulated that there might be two core systems of numerical representations – one system for representing large numerosities approximately, namely the approximate number system (ANS), and one system for representing small numbers of objects exactly. However, it has been argued that these mainly domain-specific skills are insufficient for understanding mathematical development. Thus, the contribution of both domain-general and domain-specific factors must be considered. Furthermore, a combination of domain-general abilities and domain-specific skills has been found to contribute to the growth of academic competencies, but their relative importance is not fully understood (Ferrer & McArdle, 2004; Gustafsson & Undheim, 1992; Von Aster & Shalev, 2007).

2.3.1 Domain general theories

A conceptual framework that focuses more on both domain-general and domain-specific explanations of mathematical development is suggested by Geary and Hoard (2005).

These cognitive factors comprise both cognitive functions (e.g. working memory) and mental representations (e.g. arithmetic facts stored in memory). According to Geary and Hoard (2005), both numerical knowledge and general processing abilities contribute to mathematical development. The framework was based on earlier evaluations of cognitive deficits in children with mathematical learning difficulties (Geary, 1993), and the objective was to understand the performance and cognitive patterns of children with mathematical learning difficulties.

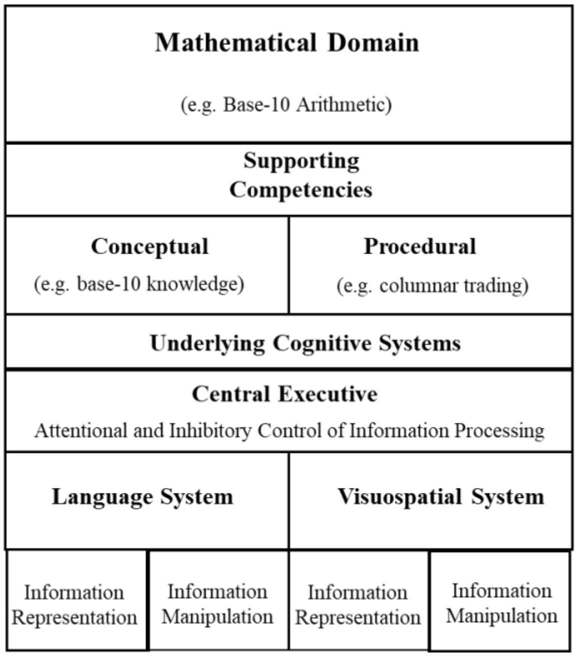

Figure 6. Hierarchical framework of the skills underpinning mathematics, adapted by Geary and Hoard (2005)

The framework depicted in Figure 6 is based on an overview of the developmental delays and deficits of children with mathematical learning difficulties, and the underpinning components of them (Geary & Hoard, 2005). The mathematical domain on the top is related to conceptual and procedural competencies, which is in turn supported by an array of cognitive systems, as shown in the bottom sections of the figure. On the basis of this framework, developmental delays and deficits in mathematics can be understood as being related to a combination of disrupted functions – the domain-general abilities such as the central executive, for instance (Geary & Hoard, 2005).

2.3.2. Working memory as a domain-general theory

In addition to Geary and Hoard's (2005) model, which perhaps can be considered as a combination of domain-general and domain-specific components, there are also other notable domain-general theories. One such domain-general theory concerns working memory. It is suggested that individual differences underlying cognitive skills predict performance in mathematics, and it has been proposed that working memory plays a substantial role in the development of numerical magnitude knowledge (Kolkman, Hoijtink, Kroesbergen, & Leseman, 2013). Working memory is found to be critical for children's procedural knowledge and for the execution of action sequences used for solving arithmetic problems (Geary, Hoard, Byrd-Craven, & De Soto, 2004). Several studies have related these abilities to both concurrent and future mathematical abilities (Bull & Lee, 2014; Cragg & Gilmore, 2014; Raghubar, Barnes, & Hecht, 2010). Although working memory has been put forward as important for mathematical achievement, the nature of this relationship is unclear (Cragg, Keeble, Richardson, Roome, & Gilmore, 2017). It is also suggested that both working memory and ANS predict performance in mathematics; hence, children displaying difficulties in both of these skills might show particularly low mathematical performance (Toll, Kroesbergen, & Van Luit, 2016).

Many studies have used Baddeley's model (Baddeley, 2000; Baddeley & Hitch, 1974) as the basis for theorizing and measuring working memory capacity. According to this model, working memory is based on different components responsible for processing visual-spatial information and verbal information. This model also includes an executive function component that organizes and plans reasoning. Geary, Hoard, and Nugent (2012) found that the role of working memory varies with both the children's experience and the type of mathematical problems they are asked to solve. Thus, the relationship between working memory and mathematics might depend on what component of working memory is focused on, but also on the type of mathematical task that is being solved. Furthermore, the correlation between mathematics and different aspects of working memory is typically diminished when domain-specific abilities are controlled for. Geary (2001) found that mathematical performance measured at Grade 1 predicted mathematical achievement and growth beyond the contribution of domain-general abilities. Fuchs and colleagues (2010) found that domain-specific skills significantly predicted simple arithmetic computational skills in Grade 1; domain-general abilities did not mediate this relation. There are mixed findings suggesting that the size of the relationship between working memory and mathematics may diminish with age. In addition, working memory is highly related to intelligence (that will be discussed in the next section); some researchers see the two as being close to being isomorphic constructs (Colom, Flores-Mendoza, & Rebello, 2003; Colom et al., 2004). Thus, typically, if intelligence is controlled for, the relationship between working memory and mathematical skills is reduced (Passolunghi, Cargnelutti, & Pastore, 2014).

2.3.3. Executive functions

It has been suggested that the key facet of working memory as related to mathematics involves attentional control and the ability to update information represented in working memory (Bull & Lee, 2014; Iuculano, Moro, & Butterworth, 2011). This suggestion opens a discussion as to whether working memory and executive functions indeed can be divided into two separate domains. Executive functions are often defined as a process that controls, directs, or coordinates other cognitive processes (Bull & Lee, 2014). Executive function skills have also been put forward as an important factor in academic success, particularly in mathematics (Allan, Hume, Allan, Farrington, & Lonigan, 2014; McClelland, Acock, & Morrison, 2006). Van der Ven, Kroesbergen, Boom, and Leseman (2011) found that updating, but not other executive functions (i.e. inhibition and shifting), explained mathematical achievement at each of four time points in Grades 1 and 2. In their study, latent growth models showed significant individual differences in the rate at which children improve in mathematics, with minimal variation in the growth rate for updating. There were significant correlations between achievements in mathematics and updating, and significant correlation in their rates of growth. Lee and Bull (2016) stated in their study, however, that there was no significant correlation between inhibition, shifting and mathematics performance, nor did they affect the correlations between working memory and mathematics.

2.3.4 Intelligence

It is clear from a number of studies that mathematics is highly related to nonverbal reasoning or what is often labelled fluid intelligence. Fluid intelligence indexes is an indication of people’s ability to identify the principal rules or concepts in new problem-solving areas (Cattell, 1963). Performance on measures of fluid intelligence and working memory are moderately correlated (Ackerman, Beier, & Boyle, 2002), but appear to assess independent competencies (Embretson, 1995; Jurden, 1995). Clearly, although there is a rather strong correlation between mathematical performance and fluid intelligence, findings from studies concerning the developmental relationship between intelligence and mathematics are contradictory. Some studies show that fluid intelligence was found to predict gains in mathematics achievement throughout childhood (e.g. Ferrer & McArdle, 2004), other find that fluid intelligence in sixth grade did not predict gains in academic skills over and beyond prior mathematical skills (Gustafsson & Undheim, 1992). This difference in results might be related to different types of analytical approaches, the use of different intelligence tests, and estimation of domain-specific effects using prior achievement. Notably, the most consistent effects are found for intelligence and the updating component of working memory (Deary, Strand, Smith, & Fernandes, 2007; Siegler et al., 2012; Östergren & Träff, 2013).

2.3.5 Domain-specific versus domain-general models

Domain-general abilities are often not able to explain development over and beyond the mathematical autoregressor. That does not imply that they are not a cause of mathematical development. They may be responsible for an early foundation, and nonverbal intelligence might be a third variable underlying both mathematics and working memory. It is suggested that longitudinal stability in mathematics persists after controlling for domain-general abilities (e.g. Aunola et al., 2004; Jordan et al., 2009). Mathematical skills are highly stable (Bodovski & Farkas, 2007), thus, things other than the autoregressor often do not come into play because there is little variation left to explain. There is still no consensus on which domain-specific or domain-general factors are fundamental to mathematical performance and mathematical development. De Smedt, Janssen, et al. (2009) found, for instance, that while executive functions predicted second grade mathematics, it was no longer a significant predictor of achievement in Grade 2 once Grade 1 mathematical achievement was included as a predictor. As with other academic domains, the relative contributions of domain-general and domain-specific abilities to subsequent mathematical achievement are not fully understood. They may vary across grades and they may vary across the level of student knowledge and mathematical content (Bailey, Watts, Littlefield, & Geary, 2014; Fuchs et al., 2016; Geary, 2011; Geary, Nicholas, Li, & Sun, 2017; Lee & Bull, 2016; Von Aster & Shalev, 2007).

2.4 What do we know and what do we not know about causal factors underpinning mathematical development?

The models and theories discussed in this section all assume causal relations in mathematical development. Strong causal claims require strong evidence (Merkley, Matejko, & Ansari, 2017). Most theories, and neither of the previously mentioned ones, have employed a rigorous randomized controlled trial design to test the causal hypothesis, and thus the results supporting the theories can be somewhat questionable. However, non-experimental work on cognitive development and in particularly longitudinal studies may be useful to generate causal hypotheses and establish the plausibility of causal relationships if they are based on clear and thoughtful interpretations of their results (Bailey, 2019). Thus, to trace possible causal influences from early cognitive skills to later arithmetic attainment, longitudinal studies starting before children enter formal education are essential (Moll et al., 2015) so that these models are tested empirically, and do not become simply box-and-arrow models. Identifying the pathways is a somewhat unique challenge. The most prominent theorists in the field of causal inference have primarily focused on estimating the effects of known causes (Shadish, Cook, & Campbell, 2002).

We do know a lot about correlation (e.g. Halberda, Feigenson, & Mazzocco, 2008) but much less about the causal relationship in mathematical development. One possible way of looking at this might be to ask whether causal mechanisms underlie the correlational evidence between domain-specific foundational competencies, for example between ANS and mathematical performance. The theoretical models mentioned do not necessarily suggest how the components of the models can be tested, but rather are just displaying the tests from which the models are derived. This raises a new challenge when it comes to drawing theoretical models. Namely, the questions about what are the reliable early and longitudinal predictors of the development of numeracy skills and other aspects of mathematics, and in what way can we strive to make the most valid inferences? Studies that attempt to disentangle the effects of different predictors and to map patterns of development are beginning to be reported (e.g. Fuchs et al., 2010; Göbel, Watson, Lervåg, & Hulme, 2014; Mazzocco, Feigeinson, & Halberda, 2011). The question is thus whether there is a two-way relationship between theoretical development and experimental testing. Recent use of latent growth modelling has started to identify the pattern of skills associated with faster growth in mathematical learning (e.g. Geary, 2005), and it would be interesting to explore how interactions here are associated with differences in learning rates.

Of the theories previously elucidated, the pathway model by LeFevre and colleagues (2010) is the only model derived and empirically tested that uses latent growth models. Perhaps theoretical models fall victim to a lack of use of more sophisticated analyses, for example the use of latent variables in structural equation modelling. This is perhaps the most relevant manner of testing theoretical models drawn from a mere hypothesis, and not deducted or tested empirically, like those of LeFevre et al. (2010).

2.5 Components of early numeracy skills

It has been suggested that humans have an innate ability to perceive, comprehend, and manipulate numerosities, a so-called number sense (Dehaene, 1997), more specifically ANS. This proposed system could address the fundamental question of where mathematical abilities stem from, and then might drive what kind of number sense we have. This core knowledge in mathematics is also found across diverse human cultures (Dehaene, Izard, Pika, & Spelke, 2006; Pica, Lerner, Izard, & Dehaene, 2004). ANS is theorized to underpin mathematical development (Feigenson et al., 2004; Piazza et al., 2010), and is usually operationalized into two skills and concepts, namely nonsymbolic and symbolic magnitude processing (De Smedt, Noël, Gilmore, & Ansari (2013). The influential and at the time innovative, and highly cited study by Halberda et al. (2008) suggested that ANS may have a causal role in determining individuals' mathematical achievement. Several subsequent studies have supported this association (Libertus, Feigenson, & Halberda, 2011; Libertus, Feigeson, & Halberda, 2013; Mazzocco et al., 2011; Schneider et al., 2016).

Thus, to examine the ANS hypothesis, we need to look at longitudinal studies and intervention studies. As for longitudinal studies, some show support for this hypothesis (Chu, vanMarle, & Geary, 2016; Halberda et al., 2008; Malone, Burgoyne, et al., 2019; Malone, Heron-Delaney, et al., 2019; Toll, Van Viersen, Kroesbergen, & Van Luit, 2015). However, these results are inconclusive. There are other longitudinal studies that do not support the ANS hypothesis (e.g. Göbel et al., 2014). Consequently, ANS is a topic of dispute. Several studies examine the association between ANS and mathematical development and seek explanations and causal patterns. In one study, Elliott et al. (2019) attempted to examine whether the relationship between ANS and mathematics is reciprocal. Using a latent cross-lagged panel model, they concluded that the relationship is indeed reciprocal, indicating that there might be a causal relationship between ANS and mathematics throughout development. The validity of the inferences concerning the causal relationship can nevertheless be challenged methodologically, for further discussion see Study 1.

2.5.1 The causal debate about the ANS

Findings suggest that preschoolers who can finely discriminate quantities have higher concurrent and later mathematics achievement than other children (Fazio, Bailey, Thompson, & Siegler, 2014; Feigenson, Libertus, & Halberda, 2013). Hence, studies have emphasized the importance of and correlations of ANS with individual mathematics achievement (e.g. Halberda et al., 2008; Mazzocco et al., 2011; Vanbinst, Ghesquière & De Smedt, 2015). These suggestions are all associational, and thus they address questions such as "what is" and not "what if" (Pearl, 2018). Pearl (2018) suggests a causal hierarchy, namely, association, intervention, and counterfactuals, association being at the bottom. This is not to say that associational studies are unimportant. They do indeed lead the way, and they are considered the first step in making causal inferences. They are not sufficient, however, for identifying the mechanisms affected by the changes that occur in the relationship between ANS and mathematical development. For obtaining causal inferences that are objective and therefore have the best chance of revealing scientific truths, carefully designed and executed randomized experiments are suggested to be optimal (Rubin, 2008).

Moreover, one could perhaps argue that ANS might not be a domain-specific skill, but a domain-general ability. If ANS is indeed a domain-general ability, meta-analyses investigating the effect of other domain-general abilities (e.g. working memory) have proven not to have a positive effect on far transfer measures (Melby-Lervåg & Hulme, 2013; Melby-Lervåg, Redick, & Hulme, 2016). These results coincide with the results of training studies of ANS. This might be due to a lack of both near transfer and far transfer in training ANS. However, currently published training studies of ANS so far have been underpowered and consequently limited (Green, Strobach, & Schubert, 2013; Moreau, Kirk, & Waldie, 2016), making interpretation difficult. Thus, ANS training studies do not show promising results (Wilson, Dehaene, Dubois, & Fayol, 2009); in addition, they display rather poor design and have methodological weaknesses. Inglis and colleagues (2017) performed a p-curve analysis of studies that reported a causal claim between ANS and mathematics performance, and argued that published and existing literature to date, do not contain enough evidence of a causal link between ANS and mathematics test.

Consequently, the question of ANS as a foundation for mathematical development remains relatively unanswered. Although, there seems to be a consensus that ANS has a role to play, the discussion is instead whether it explains anything over and beyond other numeracy skills. ANS might be seen as the first building block in early numeracy and later mathematical development, but as children learn other skills (e.g. counting), ANS might be less important in the further development. Although there are still unresolved issues related to causality in the field of mathematics, it seems clear that early numeracy skills are strongly and causally related to later mathematical achievement. The most important of these suggested causal factors are counting skills, numerical relational skills, and basic arithmetic skills. These components of early numeracy have undergone numerous investigations, with strong research designs such as longitudinal studies and randomized controlled trials. The dimensionality of these factors is unclear, given that, as will be shown, some skills that are divided into separate constructs might possibly be the same.

2.5.2 Counting skills

There seems to be support for the notion that children typically learn the count sequence by rote and then discover counting principles through informal experiences with numbers and counting (Briars & Siegler, 1984). The knowledge seems to emerge from a combination of inherent and investigational factors (Briars & Siegler, 1984; Gelman & Gallistel, 1978), and gradually young children acquire more advanced counting abilities. Gelman and Gallistel's noteworthy work from 1978 suggest that counting skills consist of five implicit principles; one-to-one correspondence; the stable order principle; the cardinality principle; the abstraction principle; and the order-irrelevance principle. Out of these, one-to-one correspondence, the stable order principle, and the cardinality principle define the "how to count" rules. Even for children who are considered poor counters, counting is described as a numerically meaningful activity (Gallistel, 2007). The one-to-one principle suggests that only one number word can be assigned to each counted object (Geary, 1994). Before kindergarten (usually around 5 years of age) most children grasp the concept that each object in a set is counted once and only once, and that the count words are always used in the same sequence for example, 1, 2, 3, 4, 5. (Gelman & Gallistel, 1978). Items must be tagged only once (Gelman & Gallistel, 1978), and children gradually learn that they can count objects presented in any configuration a long as they count each object only once (Dyson, Jordan, & Glutting, 2013). A premise for this principle is the stable order principle (Geary, 1994). The tags the child uses to match items in an array must be organized in a stable and repeated order (Gelman & Gallistel, 1978).

Coming to understand the cardinal values of the words in their count list is a lengthy process, but once achieved, it represents children's first explicit understanding of a formal mathematical concept (Carey, 2004). The achievement of learning the cardinal principle (cardinal principle knower, CPK) is suggested to be a milestone in children's mathematical achievement (Geary et al., 2018). Understanding the cardinal principle has emerged as a key predictor of later mathematical outcomes (Chu, et al., 2016; Geary & vanMarle, 2016; Geary et al., 2018). This understanding of cardinality is proposed as an early anchor for subsequent mathematics learning (Geary & vanMarle, 2016). So, what are the mechanisms behind this strong role of the cardinal principle in mathematics? Geary et al. (2017) theorized that becoming a CPK at an early age gives the CPK children more experience and practice with other mathematical skills. Consequently, children who are CPKs later have less experience being a cardinal principle knower before formal learning starts. As a result, the foundation for learning mathematics might be vulnerable. It is therefore proposed that children who understand the cardinal principle when they begin preschool (typically three- to four-years-olds) or have achieved CPK status within the first year of preschool have a substantive advantage in later number skill knowledge (Geary et al., 2018). Additionally, it is suggested that children's understanding of cardinal value of number words at the beginning of preschool predicts the sophistication of their strategy choice three years later, controlling for other factors (Geary et al., 2018).

2.5.3 Numerical relational skills

Longitudinal studies have proposed that numerical relational skills are an essential part of early numeracy development (Aunio & Niemivirta, 2010; Desoete et al., 2009; Stock et al., 2009). Numerical relational skills include a set of subskills such as the early mathematical-logical principles and understanding the meaning of the base-10 system (Aunio & Räsänen, 2016; Geary & vanMarle, 2016). Numerical relational skills also include an understanding of operational symbols in mathematics – such as more than (>), less than (<), equal to (=) and not equal to (≠) (Aunio & Räsänen, 2016). Hence, it is suggested that this componential skill include mathematical language (Negen & Sarnecka, 2012). There may be several potential reasons for why numerical relational skills are important for the development of mathematical skills. The ability to operate with number word sequences and enumerate, combined with mathematical-logical thinking is suggested to be a component of early numeracy (Aunio & Niemivirta, 2010; Desoete et al., 2009; Stock et al., 2009). Development of mathematical thinking is accordingly related to the children's growing abilities to apprehend and make relational statements, for example learning what it means that a number is equal to or more or less than another number (Aunio & Niemivirta, 2010; Resnick, 1989). The ability to numerically compare two sets is a vital aspect of the conservation ability and other related forms of numerical reasoning (e.g. Sophian, 1988).

2.5.4 Basic arithmetic skills

Primary school children are expected to learn basic arithmetic facts and learn computational procedures for solving complex arithmetic problems. Frequent and successful use of counting strategies usually leads to improvements in memory representations of arithmetical facts (e.g. counting on, use concretes or fingers as tags without counting out each addend, 4 + 3 counting, "one-two-three-four" then counting on, "five-six-seven") and leads to the strategy of retrieving arithmetical facts from long-term memory (Canobi, Reeve, & Pattison, 2002; Wilkins, Baroody, & Tilikainen, 2001). Memorization and retrieval are proposed as playing an important role in the development of arithmetic skills. Nearly all typically developing children will memorize most of the basic arithmetic facts (Geary, 2000). Arithmetic skills build on a core number knowledge system for representing numerical quantity using abstract symbols that is typically in place by the age of five (Barth, LaMont, Lipton, & Spelke, 2005). By second grade, children are typically able to answer single-digit addition problems, although rapid fact retrieval is still not mature in most children of that age (Jordan, Hanich, & Kaplan, 2003). Basic arithmetic skills in five–to-eight year-olds pertain to the degree to which a child masters mainly the addition and subtraction tasks using number symbols (Aunio & Räsänen, 2016).

2.5.5 Dimensionality

Having a closer look at both the proposed theoretical models (e.g. Aunio & Räsänen, 2016) and the components in early numeracy development (e.g. Gelman & Gallistel, 1978); there are still issues that remain to be addressed. One of these is dimensionality in some of the different components of mathematical skills. Dimensionality might be an issue when it comes to the early numeracy skills such as counting and numerical relational skills. Are there indeed two separate factors, and/or is one factor more important than the other? For instance, in the case of some of the specific skills in the core numerical skills model (Aunio & Räsänen, 2016), it can be argued that there might not be a clear-cut difference between counting and numerical relational skills. Studies conducted for instance by vanMarle, Chu, Li, and Geary (2014), describe specific skills and principles from both the relational skills, as merely elements of counting skills (e.g. ordinality). It has also been suggested that some early numeracy skills have a stronger impact on mathematical development than others. Recent studies by Geary et al. (2018) have had a stronger emphasis on one feature of counting, namely understanding the cardinal principle.

Hence, there is a need for both longitudinal studies examining both the dimensionality of mathematical development and experimental studies in order to better understand the causal relationships. There is also a need for the longitudinal studies to continue to develop and use new methodological approaches. The widely used and popular crossed-lagged panel model have for instance been subject to critique having methodological weaknesses analyzing associations between different aspects in mathematical developmental longitudinal data (Allison, 2011; Gunasekara, Richardson, Carter, & Blakely, 2014; Hamaker, Kuiper, & Grasman, 2015). For an elaboration and discussion of these panel models, see Study 1.

2.6 Mathematical learning difficulties

In the DSM-5 (American Psychiatric Association, 2013) mathematical learning difficulties are defined as *Specific Learning Disorder with impairment in mathematics.* The difficulties in mathematics must have persisted for six months, and be related to difficulties with number sense competencies, automatizing arithmetic facts, precision and accuracy related to calculations, and arithmetic reasoning. The ICD-10 (World, Health Organization, 2015), used in Norway, defines mathematical difficulties as a *Developmental arithmetical disorder*. In the Norwegian translation, it is specifically related to calculations. The definition's emphasis is on what the difficulties cannot be caused by, namely a lack of instruction and the child's cognitive abilities. Critically, mathematical skill is a continuous variable that is normally distributed in the population, and any cut-offs meant to establish normal versus disordered development will be somewhat arbitrary. However, it is common to assume that around 15–20% of children and adults experience difficulties in developing mathematical skills to such an extent that it hampers their school or work performance (Geary, 2011). Out of these, 5–7% have difficulties so severe that they are often diagnosed as having specific mathematical learning difficulties or developmental dyscalculia (American Psychiatric Association, 2013; Berch & Mazzocco, 2007; Butterworth, Varma, & Laurrilard, 2011).

Mathematical learning difficulties is suggested to be remarkably stable. According to a study by Morgan, Farkas, and Wu (2009), 65% of children who started and ended kindergarten below the tenth percentile were still in the lowest 10 percentile four years later. The average mathematics score in fifth-grade for those repeatedly displaying mathematical difficulties in kindergarten was more than two standard deviations lower than the average mathematics score in fifth grade for those who had not displayed mathematical difficulties in kindergarten (Morgan et al., 2009). Notably, children who experience poor numeracy skills in early childhood are most commonly not defined as having mathematical learning difficulties due to the young age, and if identified, they are more often called at-risk of having later mathematical learning difficulties. Even though we have longitudinal studies where children are identified with poor counting skills at the age of four, there is no tradition for labeling this before the child has started school, in Norway but also in other countries, since is recommended to evaluate the effect of intensive instruction (e.g. Geary, 2004, 2011). The scope of development is high, and it is necessary to evaluate the effect of instruction when they start formal schooling.

2.6.1 Indicators of mathematical learning difficulties

Children with mathematical learning difficulties primarily display difficulties in the skills (discussed in the previous section), which explains individual differences in mathematical achievement. This means that they have poorer performance than typically developing controls on the domain-specific skills (e.g. retrieving arithmetic facts) and domain-general skills (e.g., working memory) discussed in the previous section. There are studies indicating that children with mathematical difficulties have poorer ANS than their peers without such difficulties (Mazzocco et al., 2011; Szardenings, Kuhn, Ranger, & Holling 2018). Difficulties with basic arithmetic fact retrievals are considered a trademark for children with mathematical difficulties at an early school age, and consequently children with mathematical difficulties commit more procedural errors and use developmentally immature and error prone procedures. Moreover, many children with mathematical learning difficulties do not shift from procedural-based problem solving to memory-based solving (Ostad, 1997), suggesting that they encounter difficulties in starting or accessing arithmetic facts in or from long-term memory. One well-established finding is that typically achievers among children use a more sophisticated mix of strategies to solve simple arithmetic problems. For 7+ 4, these children either retrieve the answer, decompose the problem into easier ones (e.g. 7 + 4 = 7 + 3 = 10 + 1), or use minimal counting (starting with the larger number and counting the smaller one). Their low performing peers, in contrast, tend to commit many retrieval and counting errors, and when they count correctly, they often use immature strategies (e.g. counting both addends).

An important issue is whether the children with mathematical difficulties are qualitatively different on some variables compared with those that do not display such difficulties. Several meta-analyses have tried to address this issue (Peng & Fuchs, 2016; Peng, Namkung, Barnes, & Sun., 2016; Peng, Wang, & Namkung, 2018; Peng, Wang, Wang, & Lin, 2019). These studies have investigated the relation between domain-general abilities and mathematics. Peng et al. (2016) found that the relation between working memory and mathematics was stronger for individuals with mathematical learning difficulties. In addition, they found a significant relation between working memory and word problem-solving (Peng et al., 2016). Children with mathematical learning difficulties are also suggested to have storage and inhibition deficits specific to numerical information and dual-task deficits of both verbal and numerical information (Peng, Congying, Beilei, & Sha, 2012). This is linked to phonological storage and the phonological loop which, together with executive functions, are two key components of working memory. Phonological storage is the base of counting and simple arithmetic (Passolunghi & Siegel, 2001). The 2012 study by Peng and colleagues indicates that poor numerical storage is a distinctive feature of children with mathematical learning difficulties, furthering a notion of phonological storage deficits being a "bottleneck" that narrows information flow to the executive functions during mathematical performance (Peng et al., 2012; Peng et al., 2018).

2.6.2 Comorbidity

Several studies have confirmed the issue of comorbidity (e.g. Landerl & Moll, 2010; Raddatz, Kuhn, Holling, Moll, & Dobel, 2016). Moll et al. (2015) investigated the cognitive profiles in children with reading difficulties, mathematical difficulties, as well as children with both reading and mathematical difficulties. The results indicated that underlying cognitive deficits are not similar; children with reading difficulties have a phonological deficit, and children with mathematical learning deficits in processing numerosities (Moll et al., 2015). In a recent meta-analysis, Joyner and Wagner (2020) examined the co-occurrence of mathematical difficulties and reading difficulties, by hypothesizing mathematic difficulties as a possible predictor of reading difficulties. The main finding of this study was that children with mathematical learning difficulties are possibly two times likely to also have reading difficulties than those without mathematical learning difficulties. Since cognitive traits are correlated, children with poor mathematical skills are not only in the lower end of the distribution of mathematical skills, but often also on the lower end of distributions of other skills (Toll et al., 2016). Pennington (2006) suggested a multiple deficit model, addressing the matter of how comorbidity previously has been studied (i.e. as single indicators), supported by studies suggesting that comorbidity might be the result of a complex interplay between domain-general and domain-specific skills (Landerl & Moll, 2010). In a longitudinal study, Koponen et al. (2018) supported the notion that comorbidity might be the result of a shared domain-general factor (as for instance suggested by Willcutt et al., 2013), and also for domain-specific skills. Vanbinst, Van Bergen, Ghesquière and De Smedt (2020) investigated the co-development of early reading skills and mathematical skills, and their results conveyed significant correlations between early reading and early mathematics before children have formal instructions in school. Findings from this study displayed that phonological awareness predicted early reading and also early mathematics even after controlling for age, IQ, nonsymbolic and symbolic magnitude processing and early arithmetic (Vanbinst et al., 2020), which might indicate shared cognitive foundation of both reading and mathematical domains.

Thus, those with mathematical difficulties might more often have difficulties with reading, language and/or attention.

2.7 Early numeracy assessment

This section addresses some perspectives on educational assessment and assessment practices in Norway. Children who score poorly on mathematics achievement tests at school-entry are at elevated risk of achieving poor long-term mathematics outcomes (Duncan et al., 2007; Geary, Hoard, Nugent, & Bailey, 2012). Identifying these children and subsequently offering support has the potential to reduce the risks of developing mathematical difficulties. Given the importance of early numeracy skills for later educational and occupational success, it is crucial that educators have access to research-informed and reliable assessment of early numeracy skills.

2.7.1 Early numeracy screening

There are different ways to assess children's mathematical competence; for example, assessment interview, diagnostic assessment, dynamic assessment. Conversely, some assessments are more comprehensive and multidimensional in nature (e.g. Jordan et al., 2006, 2007; Mazzocco, 2005). In early mathematics, unlike in early literacy, no core deficit suggestive of early mathematics difficulties has been identified yet (Chiappe, 2005). Rather, mathematics skills, particularly in the early years, develop as a sequence of connected concepts and skills called a learning trajectory. As such, Foegen, Jiban, and Deno (2007) indicated that a brief assessment tool that covers a broader range of content might actually be a better means by which to assess early mathematics skills in younger children than a narrowly focused measure. Screening refers to an efficient and collective process of assessing all students in a classroom. The objective for this is to enable teachers to identify quickly and accurately which children need additional evaluation and instruction. Importantly, teachers need a range of research-informed assessment tools to meet specific needs, including the ability to screen within their classrooms.

There are, as mentioned, different approaches to assessment and the tradition or national framework for assessment approaches might lead the way when it comes to assessment approaches. Within the framework of Response to Intervention (RTI) for instance, the use of an appropriate assessment is crucial to the RTI process (Riccomini & Smith, 2011). The assessment approach in RTI can be described as twofold. RTI offers an approach with universal screening of all children in order to identify the students at risk for learning difficulties, subsequently followed for those children who were identified with the screening tool to be at risk by assessment in order to monitor those students' progress and consequently to offer them targeted support.

2.7.2 Assessment practices in Norway

What is critical to screening efforts, then, is the degree to which screeners provide accurate information about students who may need additional instruction. The RTI framework has not yet been implemented in Norway. There is no other framework establishing a national default for schools when it comes to assessment and how this is related with providing support for children. In Norway there are no national guidelines when it comes to assessment beyond the national tests provided by the directorate of education, no framework to ensure that all schools have an assessment practice comparable, for instance, with the RTI framework. Yet, in Norway, students are subject to national assessments in mathematics during the first three years of schooling (Udir, 2017). The tests are provided and developed by the Norwegian Directorate of Education and Training (Udir, 2017), which are executed towards the end of each school year pursuant to national guidelines. The national tests are summative and evaluate whether a student is over or below a set cut-off score. The schools can subsequently not choose the time during the school year that they see fit to administer it. When it comes to assessing first graders in mathematics, in Norway notably, there is no other validated assessment tool other than the national test available for Norwegian schools to use, for further elaboration see Study 2.

2.7.3 Assessment tools in Norway

To date, no study has been conducted yet about the manner and/or which assessments tools Norwegian schools use to assess young children's mathematical skills, similar to what Arnesen and colleagues (2019) did with reading and social skills. Consequentially, there is no precise overview of how teachers assess children's early numeracy beyond the tests that are publicly available. The fact that for Norwegian first graders there is only one standardized assessment in mathematics (the national test), leaves perhaps not a very controversial statement, that teachers are left having to evaluate children's mathematical development in a non-validated manner. Note that this does not mean that the assessment tools available are necessarily of poor quality, they are merely not validated yet. Nevertheless, since the tools are not validated, there is no evidence of the quality of the tool, or the lack thereof. However, whether or not there are assessments tools available or not, teachers still have to and want to assess children in mathematics. Not having validated assessment tools available might thereby result in rather arbitrary assessments. Although teachers by and large are able to single out high-performing from low-performing students through informal observation, they need assessment tools to make a more fine-grained differentiation (Kilday, Kinzie, Mashburn, & Whittaker, 2012). In Norway, students are not graded in mathematics until secondary school. There is no uniform way of assessing children's mathematical achievement and as a result, teacher assessment of children's mathematical performance is likely to vary somewhat depending on the teacher. In Norway, however, there are guidelines in the curriculum listing the mathematical competencies children are expected to have every other year, but still these guidelines are rather vague and are teacher and/or school dependent (Utdanningsdirektoratet, 2018).

2.7.4 The purpose of assessing early numeracy skills

Assessment provides us with information that guides us in supporting struggling learners. Assessment tools in early numeracy skills and early mathematics should encompass basic mathematical concepts and skills. Furthermore, they need to be reliable and valid for the purposes for which they are used, in this case, identifying the students who might need an intervention (Lee, Lembke, Moore, Ginsburg, & Pappas 2012). Although there are traditions for assessing students' achievement, it is unclear whether this practice is utilized to improve students' learning and development, how and when these assessments are being conducted, the extent to which the students' benefit from them, and/or the extent to which teachers use the assessment results. Given the importance of basic numerical skills for later educational and occupational success, it is essential that educators have access to research-informed and reliable assessment of basic numerical competencies. Efficient and reliable assessment tools can serve a twofold purpose. First, they can be used broadly to identify which children need additional instruction, and second, they can be used to identify the specific aspects of mathematical knowledge in which a particular child needs further instruction (Purpura & Lonigan, 2015). Early identification and intervention promote early success and confidence in mathematics, and can prevent later failure (Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007; Gersten, Jordan, & Flojo, 2005).

2.8 Early numeracy intervention

The previous elaboration of theoretical models concerning factors that influence numeracy development has important consequences for the construction of interventions. Thus, to get effects from an intervention, it is clearly crucial to target skills that might represent causes in the development of mathematics. As this review of theoretical models and components of early numeracy shows, there was strong support for counting, the cardinal principle, numerical relational skills, and knowing Arabic numbers as these relate to mathematical abilities. Thus, these factors would also be strong candidates to target to improve children's early mathematical skills, and as we will see, several studies have tried to do this. As for ANS, the causal status is not clear, and training studies have found no conclusive evidence that specific ANS training improves symbolic arithmetic (Szücs & Myers, 2017). There is also strong support for the notion that certain aspects of language are related to specific aspects of mathematics (i.e. vocabulary with word problems, phonological processing with arithmetic fact storage and retrieval). Some studies have introduced and applied the term mathematical language as a domain-general skill in mathematics (Purpura, Day, Napoli & Hart, 2017) and language comprehension within word-problem solving (Fuchs, Fuchs, Seethaler, & Craddock, 2019). Recent and ongoing intervention studies are also attempting to improve word-problem skills (Fuchs et al., 2019; Powell, Berry, & Barnes, 2019).

As for more domain-general processes such as working memory, executive functions, and nonverbal intelligence, these are related to mathematics, and possibly also causally related. However, it seems as if these factors are more distal causes of mathematical development, since they are rarely able to explain development in mathematical skills over and beyond previous mathematics skills. However, it is still plausible that they create a foundation for mathematical skills and for instance operate as a third variable underlying both mathematical skills and counting skills, the cardinal principle, and numerical relations. However, these kinds of distal causes are less obvious candidates to target in interventions, and as we will see, the results from studies that have tried to intervene in terms of these domain-general factors (e.g. working memory) do not show promising results (Melby-Lervåg et al., 2016).

Thus, the building blocks in the design of early numeracy instruction programs typically emphasize the early numeracy components which we know are important in early mathematical development (e.g. cardinal principle). The focus on counting skills as well as arithmetic skills is often highly valued in experimental studies in both preschool and school-aged children. Even if children have at their disposal procedures that are appropriate to a particular counting context, children might not learn which of their procedures to utilize, and fail at the task (see Gelman & Greeno, 1989). The strength of the association between early and later mathematical achievement provides a compelling rationale for schools to emphasize pertaining to instructional resources in kindergarteners and school beginners who are struggling to progress in the learning of ideas, concepts and operations in relation to early numeracy skills.

In recent years, the number of meta-analyses in the field of mathematical learning difficulties interventions has increased (Chodura, Kuhn, & Holling, 2015; Dennis et al., 2016; Gersten et al., 2009; Jitendra et al., 2018; Kroesbergen & Van Luit, 2003; Monei & Pedro, 2017; Wang, Firmender, Power, & Byrnes, 2016). Chodura et al. (2015) examined effects from interventions for children with mathematical difficulties between the ages of 6–12 years, and the results based on 35 studies showed a large increase in the intervention groups compared with the control group (d = 0.8) on measures of number skills and arithmetic. In another meta-analysis, Dennis et al. (2016) found that the overall mean effect size from the 25 studies included here was only moderate (d = 0.55). Results from this meta-analysis provide an objective for further intervention studies as well as providing this support for children in schools. However, there were methodological weaknesses, for instance many studies in the meta-analysis had low power, and no randomization. For a more detailed elaboration and discussion, see Study 3.

Moreover, other intervention studies have faced two main challenges, namely fadeout and transfer. This will therefore be discussed in more depth here, in order to investigate how these challenges might be addressed so that we can plan, design, and implement interventions accordingly.

2.8.1 Fadeout effects in mathematics

A consistent finding across interventions is fadeout effects. Fadeout brings out the pattern of diminishing effects when an effective mathematics intervention has ended (Bailey, 2019). There are different hypotheses for and perspectives on fadeout effects. Commonly, they are both complex and contextual and discussion span a range of important topics from how the fadeout effects are measured to how an intervention effect might qualitatively be prolonged with, for instance, efforts from teachers facilitating the learning outcome from an effective intervention when the intervention has ended.

2.8.2 Constraining content or preexisting differences hypothesis

The constraining content hypothesis postulates that fadeout effects are due to environmental factors after an intervention ends. This hypothesis suggests that the instruction the children in the intervention receive after the intervention does not build on knowledge they received in the intervention and that might lead to fadeout. Engel, Claessens, and Finch (2013) support this, noting that kindergarten teachers report spending a lot of the mathematical instruction time on content that most of their students already know. Bailey et al. (2016) found that children in the intervention group fell behind their posttest matched peers from the control group in the following year at approximately the same rate as the control group caught up to the treatment group. The constraining content hypothesis suggests that intervention material might be targeted at a too low level so that the control group children are likely to catch up to the treatment-group-children (Bailey et al., 2016). Successively, fadeout might be the consequence of preexiting differences hypothesis. The preexisting hypothesis argues that fadeout is due to the relatively stable variances between children (Bailey et al., 2016). This might ultimately cause them to return to their previous individual achievement trajectories after an effective early intervention (Bailey et al., 2016).

It may also be that both types of hypotheses for fadeout co-occur in a study. The intervention might have taken place too late in the school year, leaving the content taught in the intervention too easy and this in turn makes it even harder to alter the rather stable discrepancy between children. Another way of looking at it might be the hypothesis that the preexisting differences make children who struggle in early numeracy more vulnerable to instructions rendering environmental factors more important. Perhaps this is crucial for children with difficulties in early numeracy, that skills taught in an intervention need to be continually reinforced after an intervention ends, i.e. that the skills targeted in early numeracy interventions are developed as a sequence of connected concepts and skills that need scaffolding post-intervention.

2.8.3 Modest transfer

The modest transfer explanation posits that, given children's early mathematics skills are necessary for their later mathematical learning, they are necessarily not sufficient to produce long-term effects on children's mathematical achievement (Bailey et al.,2014). This fadeout effect is not exclusive to early numeracy interventions but has also been found in language interventions (Rogde, Melby-Lervåg, & Lervåg, 2016). It is therefore important to be realistic when it comes to how much changing the environment some hours per week can improve skills since the remaining environment and children's genes are stable. Thus, the fadeout effects have important implications for instruction; early interventions do not imply that the children's challenges are solved, but that children who experience problems are likely to need interventions regularly so that they do not fall back into a lower developmental trajectory.

2.8.4 Trifecta skills

To accommodate the risk of fadeout effects, three important characteristics in designing an intervention have been suggested (Bailey, Duncan, Odgers, & Yu, 2017): 1) the skills must be malleable by early intervention, 2) the skills must be fundamental for academic success, and 3) the skills that are more likely to produce persistent effects are those that do not develop quickly under counterfactual conditions. These three conditions, trifecta skills, should be aimed at targeting an intervention. Knowledge of the characteristics of children at risk of persistently low mathematics achievement may allow us to target children who, under business-as-usual conditions, would take several years to learn as they would learn from a successful early mathematics intervention (e.g. Berch & Mazzocco, 2007). Three distinct processes have been suggested to sustain benefits for children (Bailey et al., 2017); namely, skill-building, foot-in-the-door skills, and sustaining environments. Albeit being distinct, these processes are related. Both skill-building and foot-in-the-door approaches are embedded in developmental cascades theory. Within this framework, effects of early skills might have a ripple effect over time and influence later skills. Taking the cardinal principle as an example, most children will probably become cardinal principle knowers during their kindergarten years. Targeting support might not provide a permanent advantage in mathematical skills, but it might allow children more experience with these skills, which in turn can provide a more solid foundation of the development of early mathematics (Geary & vanMarle, 2016). Furthermore, it is suggested that early intervention impacts can be sustained only if they are followed by environments of sufficient quality to withstand normative growth (Bailey et al., 2017). Bailey and colleagues (2017) suggested that future studies might apply a design investigating a third treatment condition; namely, designing a third condition that specifically aims to sustain and maintain the effects of the intervention.

Consequences of fadeout effects in mathematical interventions are twofold. These perspectives are important for both the researcher planning an intervention study, and obviously also for teachers and special needs teachers giving support to children with mathematical learning difficulties. The intervention in Study 3 is the first randomized controlled trial in mathematics education ever conducted in Norway. This stresses the twofold necessity: Norway needs more intervention studies at the same time as Norwegian teachers need to adapt these trifecta perspectives when they plan, provide and maintain the targeted intervention skills; the sooner the better.

2.9 Summary

As shown in this chapter, the field of mathematical development, assessment, and interventions is characterized by both a thirst and a need for knowledge. Our understanding of these topics has developed rather quickly. From Dehaene's triple code model in 1992, to several longitudinal studies providing strong support for previous, present, and perhaps future causal claims regarding mathematical development. Drawing theoretical models (e.g. Dehaene, 1992; Geary & Hoard, 2005) and conducting correlational studies can generate hypotheses, but they do not allow us to draw causal inferences. However, they were and are indeed the starting point when it comes to finding key numeracy skills for mathematical development. In science, evidence primarily comes from empirical data, reasons, and arguments. It is key to be precise when it comes to what you have and want evidence for. Does one have evidence for an association, a causal relationship, or generalization? Research design and methodological integrity can be supportive in this (Kvernbekk, 2016). Thus, it is important to evaluate the quality and extent of evidence that presupposes the assertions. This kind of understanding allows us to ask whether the researcher has the right to legitimize a claim. Consequently there has been a shift in how we investigate mathematical development on a cognitive and behavioral level, and there are now higher criteria on what grounds causal inferences can be made. This includes for instance more robust studies when it comes to both research designs and analytical approaches. Raising the methodological bar has however lead to stronger justification of claims for the components of early numeracy that are important for mathematical development. This knowledge is necessary to identify children at risk of mathematical learning difficulties. Via these early numeracy components, we find markers of mathematical difficulties and in turn, these are the components we might need to provide these children with support.

Early numeracy skills and early mathematical development can be seen as gatekeeper skills. Opened at the right point, other skills cascade in. Moreover, if a child does not grasp these skills quickly, that door might not open again.

3 Overview of studies

3.1 Summary of Study 1

Title: Analyzing the Developmental Relationship between the Approximate Number System and Early Mathematics: Different Methodological Approaches Lead to Different Results

Authors: Lopez-Pedersen, A., Lervåg, A., & Melby-Lervåg, M.

Aim of the study

The aim of this study was to examine the developmental relationship between the approximate number system and early mathematical skills. There is a debate about the directionality of this relationship and claims on both sides of the discussion lack evidence to make strong causal inferences. With two different datasets the hypothesis that the approximate number system has a potential causal influence on mathematical development, but also that the relationship goes the other way, that mathematical skills influence ANS was examined.

Method

In order to examine whether a causal relationship between ANS and mathematical skills is plausible, it is important to test whether ANS actually explains development in mathematical skills over and beyond mathematical skills at the previous time point. Two types of parameters are specified in the CLPM, cross-lagged and autoregressive measuring respectively to what extent variability at in one time point can be explained by the variability in a different measure at a preceding time point, and an estimation of the rank-order stability of individuals. In order to differentiate the between-person and the within-person differences, the data were analyzed using random intercept cross-lagged panel model (RI-CLPM). Single indicators were used on the models in both datasets.

Results

The claim that ANS has a causal influence on mathematical skills (Halberda et al., 2008) has been highly influential in the causal debate of ANS and mathematical development. Additionally, this suggestion has furthered a hypothesis that this relationship is reciprocal (Elliott el al., 2019). However, when reanalyzing Elliott et al.'s (2019) dataset by using a random-intercept cross-lagged panel model, no evidence of a reciprocal relationship was found. The cross-lagged effects in Elliott's initial study disappeared. These results were supported having analyzed a second dataset with the same methodological approach. Results did not show any evidence supporting a reciprocal relationship between ANS and mathematical development. In sum, the results of Study 1, analyzing two different datasets with RI-CLPM, do not support the hypothesis of a reciprocal relationship between ANS and early mathematics. Study 1 also displayed how using different methodological approaches can lead to different results.

3.2 Summary of Study 2

Title: Validation of the Early Numeracy Screener for First Graders

Authors: Lopez-Pedersen, A., Mononen, R., Korhonen, J., Aunio, P., & Melby-Lervåg, M.

Aim of the study

This study investigated the psychometric properties of the Early Numeracy Screener. The Early Numeracy Screener is a teacher administered, paper-and-pencil test measuring counting skills, numerical relational skills and basic arithmetic skills. Derevied from the core numerical skills model (Aunio & Räsänen, 2016) this study aimed to investigate the dimensionality of the screener - whether the Early Numeracy Screener was identified as a one-factor model or a three-factor model.

Method

In this study, 366 first graders born in 2010 were screened with the Early Numeracy Screener. Confirmatory factor analysis (CFA) was used to examine the factor structure of the test instrument. The validity and reliability were evaluated through EFPAs quality criteria (Evers, Hagemeister, & Høstmælingen, 2013; Evers, Muñiz, Høstmælingen, Sjöberg, & Bartram, 2013). Teachers administered the screener in their respective classrooms, the screeners were then collected and data was entered by research assistants. Data about each child’s score on the national assessment in numeracy was collected towards the end of the school year once all the schools had executed it.

Results

The results indicated that the Early Numeracy Screener reliably measures three distinct subskills of early numeracy (counting skills, numerical relational skills, basic arithmetic skills). Moreover, these subskills were related to the national test scores and children identified as at risk of developing mathematical learning difficulties with the Early Numeracy test were likely to perform below the 20th percentile in the national test in mathematics six months later. The Early Numeracy Screener may thus serve as an indicator of young children's performance in early numeracy. The brevity and ease of use of the Early Numeracy Screener makes it suitable for classroom instructional settings.

3.3 Summary of Study 3

Title: Improving Numeracy Skills in Low Performing First Graders: A Randomized Controlled Trial

Authors: Lopez-Pedersen, A., Mononen, R., Aunio, A., Scherer, R., & Melby-Lervåg, M.

Aim of the study

The third study is a randomized controlled trial aimed to improve low-performing children's early numeracy skills in an eight-week intervention program in Norwegian first graders. The trial intended to answer whether there are immediate effects of the intervention, and whether these effects are lasting or fade out completely or partly as the children receive less or no intervention.

Method

All children in two municipalities in Norway born in 2010 and attending first grade were invited to participate in the study. This resulted in 366 initial participants. The children were screened with a test of early numeracy skills, and 32% children were identified with the lowest scores (N = 120, 57% girls, mean age = 77 months). The children were randomly allocated to either control or intervention group, the control group followed business as usual in their respective classrooms. Children's numeracy skills were assessed at four time points: pre-intervention, immediate posttest after 8 weeks of intervention, posttest after boost phase and a six-month follow-up.

The intervention program

The intervention program was conducted by teachers and special needs teachers working at the schools. The intervention had two phases, one initial 8-week phase with a total of 24 instruction sessions (16 small-group and 8 individual ones), approximately 130 minutes each week. The second phase of the intervention started two weeks after the first intervention phase ended. This phase involved six instruction sessions, once a week, over a total of six weeks. The sessions in the second phase were a repetition of the initial eight weeks intervention regarding their content. The objective of including a repetition phase was an attempt to avoid fadeout effects. Each of the intervention teachers received all the material in advance of the intervention start, as well as training and practice in using the material.

Results

The intervention produced positive benefits (*d* = 0.20) on early numeracy (counting and numerical relational skills) learning, but these were not significant. This may have been due to many children reaching ceiling on the measures after the intervention. There were moderate and reliable effects on word problem solving (*d* = 0.41); however, the effects were reduced and faded out on the second follow-up test (after the second intervention phase of 6 weeks once a week) and at follow-up test (6 months after the intervention) compared to the immediate post-test. Overall, this randomized controlled trial showed that early numeracy skills are malleable and that frequent and long-term interventions are needed for the positive effects to last.

4 Methodological considerations and discussion

The methodological issues and discussion in this chapter relate to several topics that have not been fully covered in the three studies. First, issues concerning methodological approaches and challenges in the three studies will be discussed. This includes limitations related to validity, design issues, analysis, and possible bias. Second, ethical considerations of research with children in studies and research designs are also included.

4.1 Methodological approaches

Choosing methodological approaches when analyzing data is continuously manifold and challenging and can have consequences for the inferences drawn. All three studies have a quantitative approach. In all the papers of this thesis, structural equation modeling has been conducted. Structural equation modeling allows one to control for measurement error, and thus makes an imperative argument for performing these types of analyses. The three studies have rather different objectives and research questions and have thus required different methodological approaches.

4.2 Methodological considerations and challenges

Study 1 initially set out to merely analyze our own dataset, exploring the data with latent-trait analysis and fixed-effect models. However, when we expanded the study to include the data from Elliott et al. (2019) that used CLPM, the story of the paper changed in the sense that the study also took a methodologically conceptual approach. We attempted thus to challenge the widely used cross-lagged panel models (CLPM). A limitation of CLPM is that they assume that every person in the study varies over time around the same means, and that there are no trait-like individual differences (Allison, 2006, 2011; Hamaker et al., 2015). As a result, the lagged parameters that are obtained with the cross-lagged panel model do not represent the actual within-individual relationships over time, and this may lead to inaccurate conclusions regarding plausible causal influences (Hamaker et al., 2015). We ended up using random intercept cross-lagged panel model, RI-CLPM (Hamaker et al., 2015). The RI-CLPM is an attempt to unravel the within-person process from the stable differences between persons since the CLPM does not differentiate between these two levels which most probably exist in the data (Hamaker et al., 2015). Using RI-CLPM the within-person level is separated from the between-person level. In RI-CLPM the between-person variance is accounted for, and thus the cross-lagged effects only represent change within-person. RI-CLPM is hence designed to obtain the estimates of cross-lagged regressions parameters that truly reflect the underlying reciprocal process that takes place at the within-person level (Hamaker et al., 2015). The fact that CLPM focuses on changes in the rank-order of individuals might disguise important developmental trends, since individuals might experience changes that does not strongly affect the overall rank-order in the sample (Mund & Nestler, 2019). For further discussion, see Study 1.

4.2.1 Sample size

Sample size is an essential issue. Structural equation modeling is a large sample technique (Kline, 2016), and sample size affects the statistical power and precision of the model's parameter estimates (Brown, 2015). Sample size has been a methodological challenge throughout the thesis, especially in Study 1 and 3. This had consequences in several respects. With a sample size of 120, we could not afford to run complex models and the sample size of 120 is almost a lower threshold for allowing structural equation modelling in the first place. In Study 1, reanalyzing Elliott et al. (2019) accommodated the challenge of sample size (n = 193), and with these data, potentially more complex models could be estimated. For example, the RI-CLPM could have been conducted with multiple indicators, and not observed variable as reported in Study 1 (i.e. both our own data and those of Elliott et al., 2019). We were not able to do this with Elliott et al.'s (2019) dataset because we only had access to the sum score of the mathematical test used in this study (TEMA-3). It would be interesting to divide this into multiple indicators, making latent variables of the observed variables in the test (e.g. counting skills separated from arithmetic skills). This might have led to a more fine-grained analysis studying the associational relationship between ANS and early numeracy skills. This, in addition to the low sample size in our own data, represented a threat to construct validity. Low sample size affects the precision of the model parameter, and thus this is a limitation when it comes to the validity of the inferences from these indicators.

4.2.2 Threats to validity in a validation study

The groundwork of assessment describes the elements that are vital when developing high-quality assessments. Test theory provides principles for high-quality assessments and psychometric properties in terms of reliability, validity, and norms. Mainly, many of these theories provide several procedures for analyzing variables and were used for evaluating the quality of Study 2 in this thesis. In Study 2, a confirmatory factor analysis was conducted to examine whether the Early Numeracy Screener could identify one or three factors. Since this was a validation study of an already established hypothesis; different aspects of threats to validity was the main objective of the study in itself. The European Federation of

Psychologists' Associations (EFPA) (Evers, Hagemeister et al., 2013; Evers, Muñiz et al., 2013) emphasizes two types of validity when it comes to assessment tools– construct validity and criterion validity. Construct validity refers to whether the items represent the theoretical constructs that they are designed for. Looking more closely into the EFPA check list for evaluations of test instruments, there are some areas that would have strengthened the soundness of the Early Numeracy Screener. Test - retest reliability was not conducted. Therefore, for future studies it would be interesting to increase the validity further with a test-retest design of the screening tool and perhaps open the possibility of doing ROC-curves analysis by adding standardized mathematical measures. Furthermore, developing an assessment tool should ideally use Item Response Theory or the Rasch Model in the initial phase; this was not done since we evaluated an already developed assessment tool. In this sense, we did not have any other criterion-related measure with which to correlate the Early Numeracy Screener (e.g. SYMP test, Brankaer, Ghesquière,& De Smedt, 2017), only the national test. The challenge with the national test was that we did not have access to the dimensionalities nor the validation of the test. We requested them, and both the test developers and the Directorate of Education did indeed want to help us with this information, but alas, they were not obtainable.

4.2.3 Validity and validation

Validity refers to the amount of relevant information a test provides, and the trustworthiness of the inferences drawn from the test scores (Thorndike & Thorndike-Christ, 2014). Whether the inferences derived from an assessment's score and interpretation of its outcomes are true, depend on how comparable the inferences are across different groups. Developing and evaluating The Early Numeracy Screener reported in this thesis endeavored to accommodate different types of validity in order to reduce biased interpretations of the test scores, and hence attempting to ensure the soundness of future teacher inferences based on having conducted the Early Numeracy Screener. In order to develop assessment tools, in this case an early numeracy screener, quality markers of an assessment tool need to meet the validity standards concerning construct validity, content validity, criterion-related validity, and reliability.

4.2.4 Construct validity

Construct validity refers to whether the items measure the theoretical construct for which they are designed. The constructs of the Early Numeracy Screener and their score interpretations reported in Study 2 are based on theories of how young children learn and develop early numeracy skills (Aunio & Räsänen, 2016). The validation of these constructs aims to confirm that they reflect the intended theoretical constructs of the assessments (Cronbach & Meehl, 1955) and a central aspect of construct validity includes evidence of content relevance (Messick, 1995). Construct validity was examined in Study 2 with confirmatory factor analysis (CFA). Furthermore, Thorndike and Thorndike-Christ (2014) suggest that construct-related evidence comes from the correspondence between the test scores to the deduction from theory. Investigating this correspondence can also be done by CFA; in Study 2, the Early Numeracy Screener was deduced from the core numerical skills model (Aunio & Räsänen, 2016), and we used CFA to investigate the dimensionality of the screener.

4.2.5 Content validity

Content validity concerns whether a test contains appropriate content and requires appropriate processes (Thornike & Thorndike-Christ, 2014). In accordance with Messick (1995), content validity alone cannot be used to qualify validity. In Study 2, the Early Numeracy Screener was developed to measure three early numeracy skills (counting skills, numerical relational skills, and basic arithmetic skills). Content validity refers to how one can tell whether the test in fact measures these skills and is related to evidence that reveals the relationship between the content and the proposed interpretation (Thorndike & Thorndike-Christ, 2014). This was accommodated using CFA to analyze the factor structure; we compared the model fit of two different models, a one-factor model and a three-factor model. The three-factor model fitted the data best and thus indicated that the three factors theorized by the core numerical skills model (Aunio & Räsänen, 2016) were identified.

4.2.6 Criterion-related validity

Criterion-related validity refers to predictive and concurrent validity, that is, whether the measurement correlates with other relevant valid assessments used for the same purpose to predict future or current performance. In Study 2, predictive validity was examined by correlating the Early Numeracy Screener scores with the national test taken later in the school year. Criterion-related validity also focuses on how norms and cut-off points are developed and how the norming sample represents the population that the assessment was designed for in terms of such characteristics as age, socio-economic background, and gender. Messick (1995) claims that the validity concepts should not be considered in isolation since they are complimentary and should therefore be viewed as a unified concept. Hence, construct validity includes aspects of content and criterion-related validity and can therefore be seen to represent validity as a whole.

4.2.7 Reliability

Both reliability and validity are required to any test. Reliability refers to the precision of a measurement procedure (Thorndike & Thorndike-Christ, 2014), scores produced by an assessment tool and how it is conducted should be consistent and reproducible. Furthermore, reliability indicates the internal consistency of a test (Thorndike & Thorndike-Christ, 2014), whether different test items that examine the same construct, (e.g. counting) produce similar result. The internal consistency of both the one-factor model and the three-factor model was satisfactory. Reliability is essential for determining the extent to which an assessment tool contains random measurement error, and this is a necessity for construct validity. In addition, the reliability of an assessment includes the accuracy of the assessment measure, and how precise the resulting scores are. The degree of reliability of a test depends on the purpose of the testing.

4.2.8 External validity

In Study 3, we investigated the effects of an intervention. With a sample size of 120, we could yet again not afford to run complex models. We did however examine the measurement models for each construct, and this process required trial and error in constructing the latent variables. For example, we tried exploring a latent variable using the scores of counting and numerical relational skills loading on one factor, but this resulted in a poor model, and did not represent the data well. Therefore, we used the two indicators of these skills as manifest scores and examined their changes over time and effects of the intervention separately.

Sample size has been a methodological challenge throughout the thesis. This had consequences in several respects. Furthermore, in Study 3, with a sample size of 120, we had to detect rather large effect sizes in order for them to be significant. The study had power to detect an effect size $d = 0.33$ in 70% of the cases with pretest as a covariate explaining 30% of the variation. One might speculate that a larger sample might have resulted in significant effects in more of the measures. The sample size, however, does affect the external validity of the intervention, namely, the validity of inferences from the context of the study to a wider context or to other contexts (Kleven, 2008). Furthermore, the low statistical power threatens external validity when it comes to whether the causal relationship holds over variations in persons, settings, treatments, and outcomes (Shadish et al., 2002). The low sample sizes in the studies in this thesis were however a result of the sparse funding of the studies. These studies had no external funding, merely departmental research funding which is quite limited, and this ultimately restricted how many participants we could include in the intervention. The total of 120 participants was thus a compromise and a prerequisite to make the study feasible.

4.3 Bias in research design – randomized controlled trials

Cartwright (2007) divides all research methods into two main groups; clinchers and vouchers. Randomized controlled trials belong to the group referred to as clinchers. Clinchers are deductive methods when specific assumptions in the experiment occur; a positive result logically follows the conclusion. The evidence is guaranteed through the research design, but if, and only if, the design gratifies all claims. Randomized controlled trials are referred to as the gold standard for causal evidence indicating "what works" (Dewey, 1929; Kvernbekk, 2013); it is a deductive method in which specific assumptions in the experiment occur. In a randomized trial, the emerging evidence is adequate for the conclusion and can be argued to be a guarantee for the conclusion (Kvernbekk, 2016). The standard result in an RCT is a treatment effect, and the objective in educational research is to improve practice. It is disputed, however, that RCTs are the gold standard in research design (Cartwright, 2007). RCTs are said to operate within relations between causes and effects, and the evidence obliges the causal claim. Cartwright and Hardie (2012) state that results from an RCT cannot be directly related to the arguments concerning effectiveness, but that results from RCTs can serve as a premise for the arguments concerning effectiveness. Cartwright (2007) argues that there are two criteria for rigorous evidence. One is credibility, the notion that credible evidence argues the likelihood that the conclusion is true. The other criterion is that explicit and complete evidence makes the conclusion likely. According to Cartwright (2007), there is no gold standard, no universally best research method, and further states that RCTs suffer from scope weakness. Cartwright (2007) argues that the results are valid for the participants in the study, but only for this precise group. The method and design do not necessarily guarantee strong claims of evidence concerning external validity.

4.4 Ethical perspectives

4.4.1 Open science framework

If this study was conducted now, the studies would have been preregistered, had open data and scripts. This would provide transparency, openness, and reproducibility in science (Nosek et al., 2015). This is important to avoid unethical research practice; however, when this thesis was begun, an open science framework had not yet become a current object of attention.

4.4.2 Children as informants

All the data from this thesis were collected from one large study; the discussion of ethical considerations will be related to children as informants and linked to the RCT study. The Norwegian Centre for Research Data (NSD) gave ethical approval, and we obtained informed parental consent for each child in this study. Parents were informed about the purpose of the study and how the results of the studies would be used as well as what participation would entail, for instance, how, when and by whom their children would be assessed, as well as what the main mechanisms of RCTs are. These include how the children are allocated to each group, what happens in the intervention group and in the control group. The information was disseminated in a neutral manner in line with the guidelines provided by the National Committees for Research Ethics in Norway (NESH) (2016). Children are considered a vulnerable group in research, and the protection of vulnerable groups is defined in the Helsinki Declaration (World Medical Association, 2013) and is thus legitimized by the objective of this study which is to ultimately provide better educational conditions for children at risk of learning difficulties in mathematics. Additionally, the researcher needs to have adequate knowledge about children to be able to adapt both methods and content of the research according to the children's age (NESH, 2016). In terms of content, the intervention was planned by researchers with backgrounds in special needs education, and thus by scholars with knowledge about children's mathematical and psychological development. In addition to the PhD student (myself), research assistants conducted the individual assessment of the 120 participants. They were master's students in special needs education and were provided with thorough instructions prior to assessing children. All the children were assessed in their respective schools, creating safe and sound testing conditions. The research assistants were instructed to take the children's well-being in the test situations into account, in research involving young children that means responsiveness to the children's needs in the test situation.

The design of randomized controlled trials also raises ethical dilemmas. The groups are not treated equally; the intervention group receives special training while the control group continues in a business-as-usual manner. An ethical dilemma exists here in the sense that it can be argued that the control group is treated unfairly (Gall, Gall, & Borg, 2007). This is also the case in Study 3, where the control group received ordinary classroom teaching.

That being said, there was not yet any evidence that this intervention would be effective; thus, we had no way of being able to predict the extent to which the intervention group might be favored. Moreover, ordinary classroom instruction is considered to be of relatively high quality in Norway; for example, there is a teacher norm stipulated by the Norwegian Directorate for Education and Training that sets a ceiling of 15 students per teacher in the early grades. The schools are obliged to comply with these guidelines.

Furthermore, we considered ethics in relation to the handling of personal data. Ethical approval and permission were obtained from the Norwegian Social Science Data Services (NSD), and anonymization, data storage, audio recordings, storage of name lists, were all ensured to comply with the requirements of NESH (2016) and NSD.

5 Discussion

The overarching topic in this thesis has been how to support children who struggle with mathematics, and thus an attempt to ascertain this through three different studies. We began with a somewhat detailed and fine-grained Study 1 in which we investigated the developmental relationship between what can be described as one of the smallest ingredients of mathematics, ANS. This study might not at first seem very relevant in terms of how we can remedy children's learning difficulties in mathematics. But if we look back at the first model presented by Morton and Frith (1995), this knowledge, involving the contesting of the causal relationship between ANS and early mathematical development, primarily represents a cognitive explanation of a learning difficulty. The model provides a theoretical understanding of the association and is line with other theoretical models for mathematical development, such as the triple code model (Dehaene, 1992) and the core numerical skills model (Aunio & Räsänen, 2016). These models have contributed to how we can support children who struggle in mathematics, because in order to do so, we need to know what early numeracy is; we need to understand how mathematical skills develop.

These theoretical models have thus furthered a fruitful line of research that has been aimed at understanding the relationship between the small elements in mathematics and the ultimate goal – proficiency in mathematics. To investigate the relationship between ANS and mathematical performance, several studies have supported the hypothesis that ANS has an effect on mathematical performance. One of the major issues in the discussion, however, is how ANS and mathematical development are causally linked. This has triggered methodological issues in this debate, and these issues were the onset of Study 1. How can different methodological approaches contribute to this discussion? When analyzing both Elliott et al.'s (2019) dataset and our own with a more valid methodological approach, ANS did not explain development in early mathematics skills over and beyond mathematical skills at previous time points. Analyzing the two datasets with RI-CLPM led to the disappearance of the reciprocal effect between ANS and mathematics postulated in Elliott et al. (2019) study, indicating that ANS is neither a cause nor an effect of mathematical development. These results weakens the support for a reciprocal relationship between ANS and mathematics. Moreover, the re-analysis of Elliott et al.'s data (2019) raises a methodological concern as different methodological approaches in the current study gave different and contrasting results.

5.1 Early numeracy assessment

The results from Study 2 indicate that the Early Numeracy Screener reliably measures three distinct subskills of early numeracy: counting skills, numerical relational skills and basic arithmetic skills. Concerning construct validity, we examined whether the Early Numeracy Screener worked best as a one-factor model or a three-factor model. The CFAs showed better model fit for the three-factor model based on counting skills, numerical relational skills and basic arithmetic skills. This supports Aunio and Räsänen's (2016) theoretical model, which suggests that the early numeracy skills is multi-dimensional. Support for assessing these distinct factors of numeracy skills is not only found in Aunio and Räsänen's (2016) study alone; but is also in previous research (Desoete et al., 2012; Gersten et al. 2012; Moeller et al., 2011). Other studies have also argued early numeracy consisting of multi factors (e.g. Cirino, 2011; Hirsch et al., 2018). The mutual challenge for these multi factor models, such as for instance Hirsch et al. (2018) and the core numerical skills model (Aunio & Räsänen, 2016) is that the model fit for these models compared to more parsimonious models are arguably rather similar. Study 2 displayed that although the three-factor model had a better fit than the one-factor model, these might indeed be highly related skills.

Given that the Early Numeracy Screener worked better as a three-factor model, this may also be beneficial for the practice field when it comes to identifying children at risk of developing mathematical difficulties. The screener enables teachers to identify and differentiate between components of early numeracy skills. Consequently, to recognize which particular skill might be difficult for a child, hence the support can be targeted and fine grained, and it might make it easier to monitor children's progress. For instance, verbal counting plays an important role as a predictor of arithmetic skills (Zhang et al., 2014); accordingly counting may be an important skill through which to identify children at risk of developing mathematical learning difficulties. Therefore, to assist children in establishing these skills, it is important to identify children who struggle at an early stage, more precisely to identify exactly which skill they are struggling with.

5.2 Intervention strength and intervention difficulty level

One of the common denominators of previously effective intervention studies (e.g. Fuchs et al., 2013; Gersten et al., 2015) is the strength of the interventions. The intervention presented in Study 3 lasted for only 8 weeks, 3 sessions per week, followed by a six week boost phase two weeks after the initial eight weeks of the intervention ended. Considering the stability in mathematics performance, the intervention in Study 3 might not have been strong enough. It should have lasted for a longer period of weeks, and the boost phase should have been planned differently. It would be interesting, for example, to have the boost phase at a later stage, perhaps a month or two after the intervention ended, and to include more sessions per week in the boost phase. Furthermore, it is possible that the content of the intervention was too easy – these skills were perhaps skills that children indeed eventually will learn. This could be a result of the rather high cut-off in our sample, the fact that these skills would not be counterfactual and thus counting and numeric relational skills training were not effective. In future studies, this could be adjusted for by having a lower cut-off; perhaps the level of this particular intervention might be more suitable for children with more severe difficulties. This does, however, require a closer look at individual differences in the participating children in this study. Which children had a greater effect from the intervention and these analysis could consequently be a starting point for developing new interventions.

The rather disappointing effects in this intervention leads one to reflect whether the intervention should have been conducted earlier in the school year. This intervention started at the end of January. That was mainly due to practicalities, since we needed roughly two months to test all the children before the intervention started. Counting skills are emphasized in the beginning of formal instruction, and in that context, it should also be noted that the control group in the current study also received high-quality instruction in early numeracy and arithmetic skills. In Norway, in first grade, there are a large number of teachers in each class, and this makes it possible to give differentiated teaching that also targets those who struggle. Thus, this opening for differentiated teaching is likely to be of high quality and enable the control group to make substantial gains when compared with the intervention group. This can perhaps explain why the gains in the intervention group on the near transfer measures of early numeracy skills such as counting and numeric relational skills and on arithmetic were somewhat disappointing.

The current intervention also shows smaller effects than the mean effects in the meta-analyses by Dennis et al. (2016) and Chodura et al. (2015). The reason for this could be that these meta-analyses include mostly quasi-experiments that are not randomized. It is well-known that quasi experiments often overestimate effects compared with randomized controlled trials (Shadish, Clarke & Steiner, 2008), and this could also be the case here. In addition, due to the practice of publishing mainly studies with significant effects, mean effects from meta-analyses are often overestimated by publication bias (Polanin, Tanner-Smith & Hennessy, 2016).

5.3 Educational implications

Since ANS is still debatable, we need perhaps to take a pragmatic viewpoint and consider the educational implications of studies of ANS. Measuring ANS in itself is not something Norwegian schools or clinicians traditionally do, and training studies with a rigorous design in mathematics is in itself understudied in Norway. Based on the discussion in both the theoretical elaboration previously mentioned and in Study 1, ANS might not be the most pertinent mathematical ability to target neither when it comes to assessing children nor in designing interventions. This is furthermore relevant as to what extent information from assessing ANS can provide pedagogical information. Perhaps what we assess should go hand-in-hand with the intended support? One might thus argue that once children at risk of developing mathematical difficulties are identified, we should target what we know they are struggling with (e.g. counting skills, arithmetic skills), and not – from the practitioners' viewpoint – the rather elusive ANS. Furthermore, if indeed ANS is a domain-general skill, meta-analyses of other domain-general skills, such as working memory, has yet to prove an effect on far transfer skills (Melby-Lervåg & Hulme, 2013; Melby-Lervåg et al., 2016). This might be applied to training studies of ANS, which have not shown conclusive evidence that specific ANS training improves arithmetic (Szücs & Myers, 2017). Hence, we should target skills that are in fact malleable and that it makes theoretical sense to target. Since there still are many unanswered questions concerning ANS, it might be more relevant to focus on skills we know are crucial for mathematical development, both when it comes to developmental milestones, assessing children at risk of mathematical difficulties and providing the necessary educational support.

5.4 Closing remarks and future directions

This thesis has been a contribution to the field of mathematical development and mathematical learning difficulties in Norway. As for educational contributions, Study 2 will have a direct implication for the Norwegian practice field. Hopefully the Early Numeracy Screener will enable schools to identify children in at risk of developing mathematical difficulties earlier in the school year. Early identification has received political attention in Norway recently (NOU 2019: 3), and the Ministry of Education has recommended that early assessment in first grade become mandatory (Kunnskapsdepartementet, 2019). Schools can now freely use the Early Numeracy Screener, and the intervention material will also be publically available for schools to use, free of charge. Study 2 and Study 3 are thus unique contributions to the Norwegian practice field. As displayed in this section, the field of mathematical development, assessment, and interventions is characterized as a field with both a thirst and need for knowledge. Additionally, the manner of how we try to understand these topics has developed rapidly.

The main questions raised in this thesis generated many additional questions. The objective of helping struggling learners such as those in Study 3 depends on knowledge of the smaller aspects, such as in those in Study 1, and the first step in helping struggling learners is to identify them. In this thesis, the originally envisioned study trajectory changed somewhat during the journey. Beginning with the enigmatic question "how", we discovered a need to challenge theoretical models and causal claims in particular. Causal relationships need to be handled with an utmost care, because this is indeed what for instance teachers who work with children strive for information about from us as research investigators. Hence, the power of the concept "effect" is immense. Moreover, as shown in this thesis and its studies, we need to operate on a methodological sound level to interpret our own results as well as those of others. As researchers one of our mandates is societal –research should also benefit society at large. In this case, society refers particularly to children, teachers, and clinicians who, in different ways, face learning difficulties in mathematics. Thus, we need to be sober and precise in terms of what we recommend that practitioners should emphasize in the identification, development and support of young children on their path to mathematics. Should we fail to discharge this duty, we thereby neglect the 'gatekeeper skills' that permit us to open doors at the right moment and influence young minds at a crucial stage of development.

6 References

Ackerman, P.L., Beier, M.E., & Boyle, M.D. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual abilities. *Journal of Experimental Psychology General, 131*(4),567–589.

Allan, N.P., Hume, L.E., Allan, D.M., Farrington, A.L., & Lonigan, C.J. (2014). Relations between inhibitory control and the development of academic skills in preschool and kindergarten: A meta-analysis. *Developmental Psychology, (50(10)*, 2368-2379.

Allison, P.D. (2006). *Fixed effects regression methods for longitudinal data using SAS*. Cary, N.Y.: SAS Institute Inc.

Allison, P.D. (2011). *Fixed Effects Regression Models.* Los Angeles: SAGE

American Psychiatric Association (Ed.). (2013). *Diagnostic and statistical manual of mental disorders. DSM-5* (5th ed). Washington, D.C.: American Psychiatric Association.

Ansari, D. (2008). Effects of development and enculturation of number representation in the brain. *Nature Reviews Neuroscience*, (9), 278-291.

Ansari, D., Donlan, C., Thomas, M.S.C., Ewing, S.A., Peen, T., & Karmiloff-Smith, A. (2003). What makes counting count? Verbal and visuospatial contributions to typical and atypical number development. *Journal of Experimental Child Psychology, 85(1)*, 50-62.

Alock, L., Gilmore, C., & Inglis, M. (2013). Experimental methods in mathematics education research. *Research in Mathematics Education, 15*(2), 97-99.

Arnesen, A., Braeken, J., Ogden, T., & Melby-Lervåg, M. (2019). Assessing children's social functioning and reading proficiency: A systematic review of the quality of educational assessment instruments used in Norwegian elementary schools. *Scandinavian Journal of Educational Research, 63(3)*, 465-490.

Aunio, P., Korhonen, J., Ragpot, L., Törmänen, M., Mononen, R., & Henning, E. (2019). Multi- factorial approach to early numeracy – The effects of cognitive skills, language factors and kindergarten attendance on early numeracy performance of South African first graders. *International Journal of Educational Research, 97*, 65-76.

Aunio, P. & Niemivirta, M. (2010). Predicting children's mathematical performance in grade one by early numeracy. *Learning and Individual Differences, 20*(5), 427–435.

Aunio, P. & Räsänen, P. (2016). Core numerical skills for learning mathematics in children aged five to eight years – a working model for educators. *European Early Childhood Education Research Journal, 24(5),* 684–704.

Aunola, K., Leskinen, E., Lerkkanen, M.-P., & Nurmi, J.-E. (2004). Developmental dynamics of math performance from preschool to grade 2. *Journal of Educational Psychology*, *96*, 699-713.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4(11)*, 417-423.

Baddeley, A., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation, 8*, 47-89.

Bailey, D.H. (2019). Explanations and implications of diminishing intervention impacts across time. In D.C. Geary, D.B. Berch, & K.M. Koepke (Eds.). *Cognitive Foundations for Improving Mathematical Learning, volume 5,* (pp. 321-346). Cambridge: Academic Press.

Bailey, D.H., Duncan, G.J., Odgers, C.L., Yu, W. (2017). Persistence and Fadeout in the Impacts of Child and Adolescent Interventions, *Journal of Research on Educational Effectiveness, 10(1)*, 7-39.

Bailey, D.H., Nguyen, T., Jenkins, J.D., Domina, T., Clements, D.H., & Sarama, J.S. (2016). Fadeout in an early mathematics intervention: Constraining content or preexisting differences. *Developmental Psychology*, *52,* 1457-1469.

Bailey, D.H., Watts, T.W., Littlefield, A.K., & Geary, D. C. (2014). State and trait effects on individual differences in children's mathematical development. *Psychological Science, 25(11)*, 2017-2026.

Barth, H., La Mont, K., Lipton, J., & Spelke, E.S. (2005). Abstract number and arithmetic in preschool children. *PNAS, 102(32)*, 14116-14121.

Barner, D., Chow, K., & Yang, S. (2009). Finding one's meaning: A test of the relation between quantifiers and integers in language development. *Cognitive Psychology*, *58*, 195– 219.

Berch, D.B., & Mazzocco, M.M.M. (2007). Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities.

Bisanz, J. & Le-Fevre, J-A. (1990). Strategic and nonstrategic processing in the development of mathematical cognition. In D.J. Bjorklund (ed.), *Children's Strategies: Contemporary Views of Cognitive Development* (pp. 213-). Hillsdale, N.J.: L. Erlbaum Associates.

Bodovski, K., & Farkas, G. (2007). Mathematics growth in early elementary school: The roles of beginning knowledge, student engagement, and instruction. *The Elementary School Journal, 108(8)*, 115-130.

Brankaer, C., Ghesquière, P., & De Smedt, B. (2017). Symbolic magnitude processing in elementary school children: A group administered paper-and-pencil measure (SYMP Test). *Behav Res 49*(4), 1361–1373.

Briars, D., & Siegler, R.S. (1984). A featural analysis of preschoolers' counting knowledge. *Developmental Psychology, 20(4)*, 607–618.

Brown, T. (2015). *Confirmatory Factor Analysis for Applied Research*. New York, NY: Guilford.

Brynner, J. M. (1997). Basic skills in adolescents' occupational preparation. *The Career Development Quarterly, 45 (4),* 305-321.

Bull., R., Espy, K.A., & Wiebe, S.A. (2008). Short-term memory, working memory, andexecutive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Journal of Developmental Neuropsychology, 33*(3), 205-228.

Bull, R., & Lee, K. (2014). Executive functioning and mathematics achievement. *Child Development Perspectives, 8(1)*, 36-41.

Butterworth, B. (2005). The development of arithmetical abilities. *The Journal of Child Psychology and Psychiatry, 46*(1), 3-18.

Butterworth, B., Varma, S., & Laurillard, D. (2011). Dyscalculia: From brain to education. *Science, 332*, 1049-1053.

Canobi, K.H., Reeve, R.A., & Pattison, P.E. (2002). *Young children's understanding of addition concepts, Educational Psychology, 22*(5), 513–532.

Carey, S. (2004). Bootstrapping and the origin of concepts. *Deadalus*, 59-68.

Cartwright, N. (2007). Are RCTs the gold standard? *BioSocieties, 2, 11-20.*

Cartwright, N. & Hardie, J. (2012) Evidence-Based Policy. A practical guide to Doing it Better. Oxford: Oxford University Press.

Cattell, R.B. (1963). Theory of fluid and crystalized intelligence: A critical experiment. *Journal of Educational Psychology, 54(1)*, 1-22.

Chiappe, P. (2005). How reading research can inform mathematics difficulties: The search for the core deficit. *Journal of Learning Disabilities, 38(4)*, 313-317.

Chodura, S., Kuhn, J.-T., & Holling, H. (2015). Interventions for children with mathematical difficulties: A meta-analysis, *Zeitschrift für Psychologie*, *223*(2), 129-144.

Chu, F.W., vanMarle, K., & Geary, D.C. (2016). Predicting children's reading and mathematics achievement from early quantitative knowledge and domain-general cognitive abilities. *Frontiers in Psychology, 7,* 1-14.

Cirino, P. T. (2011). The interrelationships of mathematical precursors in kindergarten. *Journal of Experimental Child Psychology*, *108*(4), 713–733.

Clarke, B., Smolkowski, K., Baker, S.K., Fien, H., Doabler, C.T., & Chard., D.J. (2011). The impact of a comprehensive tier 1 core kindergarten program on the achievement of students at risk in mathematics. *The Elementary School Journal, 111(4)*, 561-584.

Clements, D.H., & Sarama, J. (2008). Experimental evaluation of the effects of a research-based preschool mathematics curriculum. *American Educational Research Journal, 45(2)*, 443-494.

Colom, R., Flores-Mendoza, C., & Rebollo, I. (2003). Working memory and intelligence. *Personality and Individual Differences, 34*(1), 33–39.

Colom, R., Rebello, I., Palacious, A., Juan-Espinosa, M., & Kyllonen, P. (2004). Working memory is (almost) perfectly predicted by ***g.*** *Intelligence. 32(3)*, 277-296.

Cragg, L., & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. *Trends in Neuroscience and Education, 3(2)*, 63-68.

Cragg, L., Keeble, S., Richardson, S., Roome, H.E., & Gilmore, C. (2017). Direct and indirect influences of executive functions on mathematics achievement. *Cognition, 162*, 12-26.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52(4),* 281–302

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44*, 1-42.

Dehaene, S. (1997). *The Number Sense [How the Mind Creates Mathematics]*. Oxford: Oxford University Press.

Dehaene, S. (2001). Précis of the number sense. *Mind & Language, 16(1)*, 16-36.

Dehaene, S., & Cohen, L. (1991). Two mental calculation systems: a case study of severe acalculia with preserved approximation. *Neuropsychologica, 29*(11), 1045-1074.

Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition, 1*, 83-120.

Dehaene, S., & Cohen, L. (1997). Cerebral pathways for calculation: double dissociation between route verbal and quantitative knowledge of arithmetic. *Cortex, 33(2)*, 219-250.

Dehaene, S., Izard, V., Pica, P., & Spelke, E. (2006). Core knowledge of geometry in an Amazonian indigene group. *Science, 311*, 381-384.

Deary, I.J., Strand, S., Smith, O., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence, 35(1)*, 13-21.

Den nasjonale forskningsetiskekomité for sammfunnsvitenskap og humaniora. (2016). *Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi*. Oslo: NESH.

Dennis, M.S., Sharp, E., Chovanes, J., Thomas, A., Burns, R.M, Cister, B., & Park, J. (2016). A meta-analysis of empirical research on teaching students with mathematical learning disabilities. *Learning Disabilities Research & Practice*, *31* (3), 156-168.

De Smedt, B., Janssen, R., Bouwens, K., Verschaffel, L., Boets, B., & Ghesquière, P. (2009). Working memory and individual differences in mathematics achievement: A longitudinal study from first grade to second grade. *Journal of Experimental Child Psychology. 103(2)*, 186-201-

De Smedt, B., Noël, M-P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and education, 2(2)*, 45-55.

De Smedt, B., Verschaffel, L., & Ghesquière, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *Journal of Experimental Child Psychology, 103*(4), 469-479.

Desoete, A., Ceulemans, A., De Weerdt, F., & Pieters, S. (2012). Can we predict mathematical learning disabilities from symbolic and non-symbolic comparison tasks in kindergarten? Findings from a longitudinal study. *British Journal of Educational Psychology*, *82*, 64–81.

Desoete, A., Stock., P., Schepens., A., Baeyens., D., & Rieyers, H. (2009). Classification, seriation, and counting in grades 1, 2, 3 as two-year longitudinal predictors for low achieving in numerical facility and arithmetical achievement? *Journal of Psychoeducational Assessment*, 27 (3), pp. 252–264.

Dewey, J. (1929). The Quest for Certainty: A Study for the relation of Knowledge and Action.

De Zeeuw, E.K., de Geus, E.J., & Boomsma, D. (2015). Meta-analysis of twin studies highlights the importance of genetic variation in primary school educational achievement. *Trends in Neuroscience and Education, 4(3)*, 69-76.

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

Dyson, N.I., Jordan, N.C., & Glutting, J. (2013). A number sense intervention for low-income kindergarteners at risk for mathematics difficulties. *Journal of Learning Disabilities, 46(2)*, 166-181.

Elliott, L., Feigenson, L., Halberda, J., & Libertus, M.E. (2019). Bidirectional, longitudinal associations between math ability and approximate number system precision in childhood. *Journal of Cognition and Development, 20*(1), 56-74.

Embretson, S.E. (1995). The role of working memory capacity and general control processes in intelligence. *Intelligence, 20(2)*, 169-189.

Engel, M., Claessens, A., & Finch, M.A. (2013). Teaching students what they already know? The (mis)alignment between mathematics instructional content and student knowledge in kindergarten. *Educational Evaluation and Policy Analysis, 35(2)*, 157-178.

Evers, A., Hagemeister, C., & Hostmealingen, A. (2013). *EFPA A Review Model for the description and evaluation of psychological and educational tests.* Tech. Rep. Version 4.2.(6). Brussels: European Federation of Psychology Associations.

Evers, A., Muñiz, C., Høstmælingen, A., Sjöberg, A., & Bartram, D. (2013). Assessing the quality of tests: Revision of the EFPA review model. *Psicothema, 25*(3), pp.283–291

Fazio, L.K., Bailey, D.H., Thompson, C.L., & Siegler, R.R. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement, *Journal of Experimental Child Psychology, 123,* 53-72.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Science, 8*(7), 307-314.

Feigenson, L., Libertus, M.E., & Halberda, J. (2013). Links between the intuitive sense of number and formal mathematics ability. *Child Development Perspectives, 7*(2), 74-79.

Ferrer, E. & McArdle, J.J. (2004). An experimental analysis of dynamic hypotheses about cognitive abilities and achievement from childhood to early adulthood. *Developmental Psychology, 40(6)*, 935-952.

Fodor, J.A. (1983). The Modularity of Mind. MIT Press Ltd.

Foegen, A., Jiban, C., & Deno, S. (2007). Progress monitoring in mathematics: a review of the literature. *The Journal of Special Education, 41(2)*, 121-139.

Fuchs, L.S., Fuchs, D., Seethaler, P.M., & Craddock, C. (2019).Improving language comprehension to enhance word-problem solving. *Reading and Writing Quarterly.* https://doi.org/10.1080/10573569.2019.1666760

Fuchs, L.S., Geary, D.C, Compton, D.L., Fuchs, D., Schatschneider, C., Hamlett, C.L., … Changas, P. (2013). Effects of first-grade number knowledge tutoring with contrasting forms of practice. Journal of Educational Psychology, 105, 58-77.

Fuchs, L. S., Seethaler, P. M., Sterba, S. K., Craddock, C., Fuchs, D., Compton, D. L., … Changas, P. (2019). Word-problem intervention with and without embedded language comprehension instruction: Causal evidence on language comprehension's contribution to word-problem solving. Manuscript submitted for publication.

Fuchs, L. S., Fuchs, D., Stuebing, K., Fletcher, J. M., Hamlett, C. L., & Lambert, W. (2008). Problem solving and computational skill: Are they shared or distinct aspects of mathematical cognition? *Journal of Educational Psychology, 100*(1), 30–47.

Fuchs, L.S., Geary, D.C., Compton, D.L., Fuchs, D., Hamlett, C.L., & Bryant, J.D. (2010). The contributions of numerosity and domain-general abilities to school readiness. *Child Development, 81(5),* 1520-1533.

Fuchs, L.S., Gilbert, J.K., Powell, S.R., Cirino, P.T., Fuchs, D., Hamlett, C.L., Seethaler, P.M., & Tolar, T.D. (2016). The role of cognitive processes, foundational math skill,and calculation accuracy and fluency in word-problem solving versus prealgebraic knowledge. *Developmental Psychology, 52(12)*, 2085-2098.

Gall, M., Gall, J., & Borg, R. (2007). Educational research: An introduction (8th ed.). New York, NY: Pearson Education.

Gallistel, C.R. (2007). Commentary on Le Corre & Carey. *Science Direct Cognition, 105*, 439-445.

Geary, D.C. (1993). Mathematical disabilities: cognitive, neuropsychological, and genetic components. *Psychological Bulletin, 114()*, 345-362.

Geary, D.C. (1994). *Children's mathematical development: research and practical applications*. Washington D.C: American Psychological Association.

Geary, D.C. (2000). From infancy to adulthood: the development of numerical abilities. *European Child & Adolescent Psychiatry, 9*, 11-16.

Geary, D.C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities. 37(1)*, 4-15.

Geary, D.C. (2005). Role of Cognitive Theory in the Study of Learning Disability in Mathematics. *Journal of Learning Disabilities, 38(4)*, 305-307.

Geary, D. C. (2010). Missouri longitudinal study of mathematical development and disability. *British Journal of Educational Psychology Monograph Series II, 7,* 31–49.

Geary, D.C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology, 47(6)*, 1539-1552.

Geary, D.C., Bailey, D.H., & Hoard, M.K. (2009) Predicting mathematical achievement and mathematical learning disability with a simple screening tool: The Number Set Test. *Journal of Psychoeducational Assessment, 27*(3), 265-279.

Geary, D.C., & Hoard, M.K. (2005). Learning disabilities in arithmetic and mathematics: Theoretical and empirical perspectives. In J.I.D. Campbell (Ed.), *Handbook of Mathematical Cognition* (pp.253-268). New York: Psychology Press.

Geary, D.C., Hoard, M.K., Byrd-Craven, J., Nugent, L., & Numtee. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development, 78(4)*, 1343-1359.

Geary, D.C., Hoard, M.K., Byrd-Craven, J. & De Soto, M.C. (2004). Strategy choices in simple and complex addition: Contributions from working memory and counting knowledge for children with mathematical disability. *Journal of Experimental Child Psychology, 88*, 121-151.

Geary, D.C., Hoard, M.K., & Nugent, L. (2012). Independent contributions of the central executive, intelligence, and in-class attentive behavior to developmental change in the strategies used to solve addition problems. *Journal of Experimental Psychology, 113(1)*, 49-65.

Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2012). Mathematical cognition deficits in children with learning disabilities and persistent low achievement: A five-year prospective study. *Journal of Educational Psychology, 104*(1), 206–223.

Geary, D. C., Nicholas, A., Li, Y., & Sun, J. (2017). Developmental change in the influence of domain-general abilities and domain-specific knowledge on mathematics achievement: An eight-year longitudinal study. *Journal of Educational Psychology, 109*(5), 680–693.

Geary, D.C. & vanMarle, K., (2016). Young children's core symbolic and nonsymbolic quantitative knowledge in the prediction of later mathematics achievement. *Developmental Psychology, 52*(12), 2130–2144.

Geary, D.C., vanMarle, K., Chu, F.W, Rouder, J., Hoard, M.K., & Nugent, L. (2018). Early Conceptual Understanding of Cardinality Predicts Superior School-Entry Number-System Knowledge. *Psychological Science, 29(2)*, 191-205.

Gelman, R., & Gallistel, C.R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

Gelman, R., & Greeno, J. G. (1989). On the nature of competence: Principles for understanding in a domain. In L. B. Resnick (Ed.)**,** *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 125–186). Lawrence Erlbaum Associates, Inc.

Gersten, R., Clarke. B., Jordan, N.C., Newman-Gonchar, R., Haymond, K., & Wilkins, C. (2012). Universal screening in mathematics for the primary grades: Beginnings of a research base. *Exceptional Children*, *78*, 423–445.

Gersten, R., Chard, D.J., Jayanthi, M., Baker, S.K., Morphy, P., & Flojo, J. (2009). Mathematics Instruction for Students with Learning Difficulties: A meta-analysis of instructional components. *Review of Educational Research*, *79*, 1202-1242.

Gersten, R., Jordan, N.C., & Flojo, J.R. (2005). Early identification and interventions for students with math difficulties. *Journal of Learning Difficulties*, *38*, *293–304.*

Gersten, R., Rolfhus, E., Clarke, B., Decker, L., Wilkins, C., & Dimino, J. (2015). Intervention for first graders with limited number knowledge: Large-scale replication of a randomized controlled trial. *American Educational Research Journal*, *52*, 516-546.

Green, C.S., Strobach, T., & Schubert, T. (2013). On methodological standards in training and transfer experiments. *Psychological Research, 78*, 756-772.

Gunasekara, F.I., Richardson, K., Carter, K., & Blakely, T. (2014). Fixed effects analysis of repeated measures data. *International Journal of Epidemiology, 43*(1), 264-269.

Gustafsson, J-E., & Undheim, J.O. (1992). Stability and change in broad and narrow factors of intelligence from ages 12 to 15 years. *Journal of Educational Psychology, 84(2)*, 141-149.

Göbel, S.M., Watson, S.E., Lervåg, A., & Hulme, C. (2014). Children's arithmetic development: It is number knowledge, not the approximate number sense, that counts. *Psychological science, 25*(3), 798-798.

Halberda, J., Mazzocco, M.M.M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature, 455,* 665-668 doi:10.1038/nature07246

Hamaker, E.L., Kuiper, R.M., & Grasman, R.P.P.P. (2015). A critique of the cross-lagged panel model, *Psychological Methods, 20*(1), 102-116.

Hannula-Sormunen, M.M., Lehtinen, E., & Räsänen, P. (2015) Preschool Children's Spontaneous Focusing on Numerosity, Subitizing, and Counting Skills as Predictors of Their Mathematical Performance Seven Years Later at School. *Mathematical Thinking and Learning, 17(2-3),* 155-177.

Hart, S. A., Petrill, S. A., Thompson, L. A., & Plomin, R. (2009). The ABCs of math: A genetic analysis of mathematics and its links with reading ability and general cognitive ability. *Journal of Educational Psychology, 101*(2), 388–402.

Hirsch, S., Lambert, K., Coppens, K., & Moeller, K. (2018). Basic numerical competences in large-scale assessment data: Structure and long-term relevance. *Journal of Experimental Child Psychology*, *167*, 32–48.

Holloway, I.D., & Ansari, D. (2009). Mapping numerical magnitude onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology, 103(1),* 17-29.

Howell., S.C., & Kemp., C.R. (2010). Assessing preschool number sense: skills demonstrated by children prior to school entry. *Journal of Educational Psychology , 30*(4), 411-429.

Hulme, C., & Snowling, M. (2009). *Developmental disorders of language learning and cognition*. Chichester: Wiley-Blackwell.

Inglis, M., Batchelor, S., Gilmore, C., Watson, D. (2017). *The Behavioral and Brain Sciences*, *40*, e174. https://doi.org/10.1017/S0140525X16002120

Iuculano, T., Moro, R., & Butterworth, B. (2011). Updating working memory and arithmetical attainment in school. *Learning and Individual Differences, 21(6)*, 655-661.

Jitendra, A.K., Lein, A.E., Im, S.-h., Alghamdi, A.A., Hefte, S.B., & Mouanoutoua, J. (2018) Mathematical interventions for secondary students with learning disabilities and mathematics difficulties: A meta-analysis. *Exceptional Children*, *84*(2), 177-196.

Jordan, N.C., Glutting, J., Dyson, N., Hassinger-Das, B., & Irwin, C. (2012). Building kindergartners' number sense: A randomized controlled study. *Journal of Educational Psychology*, *104*, 647–660.

Jordan, N.C., Glutting, J., & Ramineni, C. (2008). A number sense screening tool for young children at risk for mathematical difficulties. In A. Dowker (Ed.), *Mathematical Difficulties: Psychology, neuroscience and intervention* 8pp. 45-58). New York: Elsevier.

Jordan, N.C., Hanich, L.B., & Kaplan, D. (2003). A longitudinal study of mathematical competencies in children with specific mathematics difficulties versus children with comorbid mathematics and reading difficulties. *Child Development, 74(3)*, 834-850.

Jordan, N.C., Kaplan, D., Locuniak, M.N., & Ramineni, C. (2007). Predicting first-grade math achievement from developmental number sense trajectories. Learning Disabilities, Research & Practice, 22(1), 36-46.

Jordan, N.C., Kaplan, D., Olàh, L.N., & Locuniak, M.N. (2006), Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematical difficulties. *Child Development, 77(1)*, 153-175.

Jordan, N.C., Kaplan, D., Ramineni, C., & Locuniak, M.N. (2009). Early math matters: Kindergarten number competence and later mathematics outcome. *Developmental Psychology*, *45*, 850-867.

Joyner, R.E. & Wagner, R.E. (2020). Co-occurrence of reading disabilities and math disabilities: a meta-analysis. *Scientific studies of reading, 24(1), 14-22.*

Jurden, F.H. (1995). Individual differences in working memory and complex cognition. *Journal of Educational Psychology, 8(1)*, 93-102.

Kilday, C.R., Kinzie, M.B., Mashburn, A.J., & Whittaker, J.S. (2012). Accuracy of teacher judgements of preschoolers' math skills. *Journal of Psychoeducational Assessment, 30(2)*, 148-159.

Klein, A., Starkey, P., Clements, D., Sarama, J., & Iyer, R. (2008). Effects of a Pre-Kindergarten Mathematics Intervention: A Randomized Experiment. *Journal of Research on Educational Effectiveness, 1(3).* 155-178.

Kleven, T.A. (2008). Validity and validation in qualitative and quantitative research. *Nordic Studies in Education, 28*, 21-233.

Kline, R.B. (2016). *Principles and practice of structural equation modeling.* New York: The Guilford Press.

Kolkman, M.E., Hoijtink, H.J.A., Kroesbergen, E.H., & Leseman, P.P.M. (2013). The role of executive functions in numerical magnitude skills. *Learning and Individual Differences,24*, 145-151.

Koponen, T, Aro, M., Poikkeus, A-M., Niemi, P., Lerkkanen, M-K., Ahonen, T., & Nurmi, J-E. (2018). Comorbid fluency difficulties in reading and math: Longitudinal stability across early

Krajewski, K. & Schneider, W. (2009). Exploring the impact of phonological awareness, visual- spatial working memory, and preschool quantity-number competencies on mathematics achievement in elementary school: Findings from a 3-year longitudinal study. *Journal of Experimental Child Psychology*, *103*, 516-531.

Kroesbergen, E.H. & Van Luit, J.E.H. (2003). Mathematics interventions for children with special educational needs: A Meta-Analysis. *Remedial and Special Education*, *24*, 97-114.

Kunnskapsdepartementet (2019-2020). *Tett på – tidlig innstas og inkluderende felleskap i barnehage, skole og SFO. Retreived from:* https://www.regjeringen.no/no/dokumenter/meld.-st.-6-20192020/id2677025/ [Ministry of Education and Research]

Kvernbekk, T. (2013). Philosophy of Science. A Brief Introduction to Selected Topics: Categorization, Justification, and the Relation between Observation and Theory. I:

B.H. Johnsen (red.), *Research Project Preparation within Education and Special Needs Education. (p. 35-59).* Oslo: Cappelen Damm AS.

Kvernbekk, T (2016). Evidence-based Practice in Education. Functions of evidence and causal presuppositions. London: Routledge.

Landerl, K. & Moll, K. (2010). Comorbidity of learning disorders: prevalence and familial transmission. *The Journal of Child Psychology and Psychiatry, 51(3)*, 287-294.

Lee, K., & Bull, R. (2016). Developmental changes in working memory, updating, and math achievement. *Journal of Educational Psychology, 108(6)*, 869-882.

Lee, Y-S., Lembke, E., Moore, D., Ginsburg, H.P., & Pappas, S. (2012). Item-level and construct evaluation of early numeracy curriculum-based measures. *Assessment of Effective Intervention, 37(2)*, 107-117.

LeFevre, J-A., Fast, L., Skwarchuk, S-L., Smith-Chant, B.L, Bisanz, J., Kamawar, D., & Penner-Wilger, M. (2010). Pathways to mathematics: Longitudinal predictors of performance. *Child Development, 81(6)*, 1753-1767.

Libertus, M.E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science, 14*, 1292-1300.

Libertus, M.E., Feigenson, L., & Halberda, J. (2013). Numerical approximation abilities correlate with and predict informal but not formal mathematics abilities. *Journal of Experimental Child Psychology, 166* (4), 829-838.

Lira, C.J., Carver, M., Douglas, H., & LeFevre, J-A. (2017). The integration of symbolic and non-symbolic representations of exact quantity in preschool children. *Cognition, 166*, 382-397.

Malone, S.A., Burgoyne, K., & Hulme, C. (2019). Number knowledge and the approximate number system are two critical foundations for early arithmetic development. *Journal of Educational Psychology.* Advance online publication. http://dx.doi.org/10.1037/edu0000426

Malone, S.A., Heron-Delaney, M., Burgoyne, K., & Hulme, C. (2019). Learning correspondences between magnitudes, symbols and words: Evidence for a triple code model of arithmetic development. *Cognition. 187*, 1-9.

Mazzocco, M.M., (2005). Challenges in identifying target skills for math disability screening and intervention. *Journal of Learning Disabilities, 38(4)*, 318-323.

Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011). Impaired acuity of the approximate number system underlies mathematical learning disability (dyscalculia). *Child Development, 82* (4), 1224-1237.

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. *Early Childhood Research Quarterly, 21(4)*, 471-490.

Melby-Lervåg, M., Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49*, 270–291.

Melby-Lervåg,M., Redick, T.S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of „far transfer“: evidence from a meta-analytic review. *Perspectives on Psychological Science, 11(4)*, 512-534.

Merkley, R., Matejko, A.A., & Ansari, D. (2017). Strong causal claims require strong evidence: A commentary on Wang and colleagues. *Journal of Experimental Child Psychology, 153*, 163-167.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, Vol 50*(9), 741–749.

Moeller, K., Pixner, J., Zuber, J., Kaugmann, L. & Nuerk, H.-C. (2011). Early place-value understanding as a precursor for later arithmetic performance - A longitudinal study on numerical development. *Research in Developmental Disabilities*, *38*, 1837–1851.

Moll, K., Snowling, M.J., Göbel, S.M., & Hulme, C. (2015). Early language and executive skills predict variations in number and arithmetic skills in children at family-risk of dyslexia and typically developing controls. *Learning and Instruction, 38*, 53-62.

Monei, T. & Pedro, A. (2017). A systematic review of interventions for children presenting with dyscalculia in primary schools. *Educational Psychology in Practice*, *33* (3), 277-293.

Moreau, D., Kirk, I.J., & Waldie. K.E. (2016). Seven pervasive statistical flaws in cognitive training interventions. *Frontiers in Human Neuroscience, 10*, 1-17.

Morgan, P-L-, Farkas, G., & Wu, Q. (2009). Kindergarten Predictors of Recurring Externalizing and Internalizing Psychopathology in the Third and Fifth Grades. *Journal of Emotional and Behavioral Disorders, 17(2)*, 67-79.

Morton, J., & Frith, U. (1995). Causal modeling: A structural approach to developmental psychopathology. In D. Cicchetti & D.J. Cohen (Eds.), *Manual of developmental psychopathology* (pp. 357-390). New York: Wiley.

Muldoon, K., Towse, J., Simms, V., Perra, O., & Menzies, V. (2013). A longitudinal analysis of estimation, counting skills, and mathematical ability across the first school year**.** *Developmental Psychology, 49*(2), 250-257.

Muthèn, L. K., and O. Muthèn. (1998–2017). Mplus user’s guide. 3rd ed. Los Angeles, CA: Muthèn & Muthèn.

Negen, J., & Sarnecka, B.W. (2012). Number-concept acquisition and general vocabulary development. *Child Development, 83(6)*, 2019-2027.

Nosek, B.A., Alter, G., Banks, D., Borsbroom, S.D., Dowman, S.D., Breckler, S.J. …Yorkoni, T. (2015). Promoting an open research culture. *Science, 348(6242)*, 1422-1425.

NOU 2019:3 *Nye sjanser – bedre læring – Kjønnsforskjeller i skoleprestasjoner og utdanningsløp*. Oslo: Kunnskapsdepartementet.

Ostad, S.A. (1997). Developmental differences in addition strategies: a comparison of mathematically disabled and mathematically normal children. *British Journal of Educational Psychology, 67(3)*, 345-357.

Ostad, S.A. (1998). Developmental differences in solving simple arithmetic word problems and simple number-fact problems: A comparison of mathematically normal and mathematically disabled children. *Mathematical Cognition, 4(1)*, 1-19.

Ostad, S.A. (2000). Cognitive subtraction in a developmental perspective: Accuracy, speed-of-processing and strategy-use differences in normal and mathematically disabled children. *Focus on Learning Problems in Mathematics Spring Edition, 22(2)*, 18-31.

Ostad, S.A. & Sorensen, P.M. (2007). Private speech and strategy-use patterns: bidirectional comparisons of children with and without mathematical difficulties in a developmental perspective. *Journal of Learning Disabilities, 40(4)*, 2-14.

Passolunghi, M.C., Cargnelutti, E., & Pastore, M. (2014). The contribution of general cognitive abilities and approximate number system to early mathematics. *British Journal of Educational Psychology, 84(4)*, 631-649.

Passolunghi, M.S., & Siegel, L.S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *Journal of Experimental Child Psychology, 80(1)*, 44-57.

Peng, P., Congying, S., Beilei, L., & Sha, T. (2012). Phonological storage and executive function deficits in children with mathematics difficulties. *Journal of Experimental Psychology, 112(4)*, 452-466.

Peng, P., & Fuchs, D. (2016). A Meta-Analysis of Working Memory Deficits in Children with Learning Difficulties: Is There a Difference Between Verbal Domain and Numerical Domain? *Journal of Learning Disabilities, 49(1)*, 3-20.

Peng, P., Namkung, J., Barnes, M., & Sun, C. (2016). A meta-analysis of mathematics and working memory: Moderating effects of working memory domain, type of mathematics skill, and sample characteristics. *Journal of Educational Psychology, 108(4),* 455-473.

Peng, P., Wang, C., & Namkung, J. (2018). Understanding the cognition related to mathematics difficulties: a meta-analysis on the cognitive deficit problems and the bottleneck theory. *Review of Educational Research, 88(3)*, 434-476.

Peng, P., Wang, T., Wang, C., & Lin, X. (2019). A meta-analysis on the relation between fluid intelligence and reading/mathematics: Effects of tasks, age, and social economic status. *Psychological Bulletin, 145(2),* 189–236.

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D. ... Zorzi, M. (2010). Developmental trajectory of number acuity reveals a severe impairment in develop-mental dyscalculia. *Cognition*, 116(1), 33–41.

Pica, P., Lerner, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science, 306*, 499-503.

Pearl, J. (2018). Theoretical impediments to machine learning with seven sparks from the causal revolution. *arXiv:1801.04016v1, 1-8.*

Polanin, J. R., Tanner-Smith, E. E., & Hennessy, E. A. (2016). Estimating the difference between published and unpublished effect sizes: A meta-review. *Review of Educational Research*, *86*(1), 207-236.

Powell, S. R., Berry, K. A., & Barnes, M. A. (2019). Effects of a word-problem intervention on word-problem language features for third-grade students with mathematics difficulty. Learning Disabilities, 24(2).

Praet, M., & Desoete, A. (2014). Number line estimation from kindergarten to grade 2: A longitudinal study. *Learning and Instruction, (33)*, 19-28.

Praet, M., Titeca, D., Ceulemans, A., & Desoete, A. (2013). Language in the prediction of arithmetics in kindergarten and grade 1. *Learning and Individual Differences, (27)*, 90-96.

Purpura, D.J., Baroody, A.J., & Lonigan, C.J. (2013). The transistion from informal to formal mathematical knowledge: Mediation by numeral knowledge. *Journal of Educational Psychology, 105(2)*453-464.

Purpura, D.J., Day, E., Napoli, A.R., & Hart, S.A. (2017). Identifying domain-general and domain-specific predictors of low mathematics performance: A classification and regression tree analysis. *Journal of Numerical Cognition, 3(2)*, 365-399.

Purpura, D.J., Hume, L.E., Sims, D.M., & Lonigan, C.J. (2011). Early literacy and early numeracy: The value of including early literacy skills in the prediction of numeracy development, *110(4)*, 647-658.

Purpura, D.J., & Lonigan, C.J., (2013). Informal numeracy skills: the structure and relations among numbering, relations and arithmetic operations in preschool. American *Educational Research Journal* 50(19, pp. 178–209.

Purpura, D.C., & Lonigan, C.J. (2015) Early Numeracy Assessment: The Development of the Preschool Early Numeracy Scales, Early Education and Development, 26(2), 286-313.

Raddatz, J., Kuhn, J-T., Holling, H., Moll, K., & Dobel., C. (2016). Comorbidity of arithmetic and reading disorders: basic number processing and calculation in children with learning mathematics. *Journal of Learning Disabilities, 50(3)*, 298-308.

Raghubar, K.P., Barnes, M.A., & Hecht, S.A. (2010). Working memory and mathematics: A reviewof developmental, individual difference, and cognitive approaches. *Learningand Individual Differences, 20(2)*, 110-122.

Ramani, G.B., & Siegler, R.S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development,79(2)*, 375-394.

Reeve, R.A., Paul. J.M., & Butterworth, B. (2015). Longitudinal changes in young children's 0-100 to 0-1000 number-line errors signatures. *Frontiers in Psychology (6)647*, 1-9.

Resnick, L.B. (1989). Developing mathematical knowledge. *American Psychologist, 44(2),* 162–169.

Riccomini, P.J., & Smith, G.W. (2011). Introduction to response to intervention in mathematics. In R.Gersten & R. Newman-Gonchar (Eds.). *Understanding RTI in Mathematics* (pp.1-16). Los Angeles: Paul H. Brookes.

Ritchie, S.J. & Bates., T.C. (2013). Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. *Psychological Science, 24*(7), 1301-1308.

Rittle-Johnson, B., & Siegler, R.S. (1998) The relation between conceptual and procedural knowledge in learning mathematics: A review. In C. Donlan (Ed.), *The development of mathematical skill* (pp. 75-110). Hove, England: Psychology Press.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*(2), 346-362.

Rivera-Batiz., F.L. (1992). Quantitative literacy and the likelihood of employment among young adults in the United States. *The Journal of Human Resources, 27*(2), 313-328.

Rogde, K., Melby-Lervåg, M., & Lervåg, A. (2016). Improving general language skills in second-language learners in kindergarten: A randomized controlled trial. *Journal of Educational Effectiveness*, 9(1), 155-170.

Rubin, D. B. (2008). For objective causal inference, design trumps analysis. *The Annals of Applied Statistics, 2(3)*, 808-840.

Salminen, J.B., Koponen, T.K., & Tolvanen, A.L. (2018). Individuality in the early number skill components underlying basic arithmetic skills. *Frontiers in Psychology, 9*, pp. 1-11.

Sarnecka, B.W., Kamenskaya, V.G., Yamana, Y., Ogura, T., & Yudovina, Y.B. (2007). From grammatical number to exact numbers: Early meanings of 'one', 'two', and 'three' in English, Russian, and Japanese. *Cognitive Psychology, 55(2)*, 136-168.

Schneider, M., Beeres, K., Coban, L., Merz, S., Schmidt, S.S., Strickler, J., & De Smedt, B. (2016). Associations on non-symbolic and symbolig numerical magnitude processing with mathematical competence: a meta-analysis, *Developmental science, 20*(3), pp.1-16.

Shadish, R., Cook, T.D., & Campbell, D. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston: Houghton Mifflin Company.

Shadish, W. R., Clark, M. H., & Steiner, P. M. (2008). Can nonrandomized experiments yield accurate answers? A randomized experiment comparing random and nonrandom assignments. *Journal of the American Statistical Association*, *103*(484), 1334-1344.

Siegler, R.S. (1988). Individual differences in strategy choices: good students, not-so-good students, and perfectionists. *Child Development, 59(4)*, 833-851.

Siegler, R.S. **(**2016**).** Magnitude knowledge: the common core of numerical development. *Developmental Science, 19(3),*341–61.

Siegler, R.S. & Booth, J.L. (2005). Development of Numerical estimation. In J.I.D. Campbell (Ed.), *Handbook of Mathematical Cognition* (pp. 197-212). New York: Psychology Press.

Siegler, R.S., Duncan, G.J., Davis-Kean, P.E., Duckworth, K., Claessens, A., Engel, Susperreguy, M.I., & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science, 23(7)*, 691-697.

Siegler, R.S. & Ramani, G.B. (2008). Playing linear numerical board games promotes low-income children's numerical development. *Developmental Science, 11(5)*, 655-661.

Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games—but not circular ones—improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology, 101(3*), 545–560.

Siemann, J., & Petermann, F. (2018). Evaluation of the triple code model of numerical processing – reviewing past neuroimaging and clinical findings, *72*, 106-117.

Sophian, C. (1988). Limitations on preschool children's knowledge about counting: Using counting to compare two sets. *Developmental Psychology, 24*(5), 634–640.

Sowinski, C., LeFevre, J-A., Skwarchuk, S-L., Kamawar, D., Bisanz, J., & Smith-Chant, B. (2015). Refining the quantitative pathway of the Pathways to Mathematics model. *Journal of Experimental Child Psychology, 131*, 73-93.

Stock, P., Desoete, A., & Roeyers, H. (2009). Detecting children with arithmetic disabilities from kindergarten: Evidence from a 3-year longitudinal study on the role of preparatory arithmetic abilities. *Journal of Learning Disabilities*, *43*, 250–268.

Storch, A.S., & Whitehurst, G.J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38(6)*, 934-947.

Swanson, H.L., Jerman, O., & Zheng, X. (2008). Growth in working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology, 100(2)*, 343-379.

Szardenings, C., Kuhn, J-T., Ranger, J., & Holling, H. (2018). A diffusion model analysis of magnitude comparison in children with and without dyscalculia: care of response and ability related to both mathematical achievement and stimuli. *Frontiers in Psychology, 8*, 1-11.

Szücs, D., & Myers, T. (2017). A critical analysis of design, facts, bias and inference in the approximate number system training literature: A systematic review. *Trends in Neuroscience and Education*, 1–37.

Thorndike, R.M., & Thorndike-Christ.T. (2014). *Measurement and evaluation in psychology and education* (8th ed. Ed.). Harlow: Pearson.

Trilling, B. & Fadel, B. (2009). *21st century skills:learning for life in our times*. San Fransisco: Jossey-Bass.

Toll, S.W.M., Kroesbergen, E., & Van Luit, J.E.H. (2016). Visual working memory and number sense: Testing the double deficit hypothesis in mathematics. *British Journal of Educational Psychology, 86(3)*, 429-445.

Toll, S.W.S., Van Viersen, S., Kroesbergen, E.H., Van Luit, J. (2015). The development of (non-)symbolic comparison skills throughout kindergarten and their relations with basic mathematical skills. *Learning and Individual Differences, 38,* 10-17.

Utdanningsdirektoratet (2018). *Lærernorm for grunnskolenkalkulator (2019-2010)*. Retrieved from www.udir.no/tall-og-forskning/statistikk/statistikk-runnskole/larernormkalkulator/larernorm-for-grunnskolen2/ [The Norwegian directorate of education and training]

Utdanningsdirektoratet (2017). Nasjonale kartleggingsprøver. Hentet fra https://udir.no/eksamen-og-prover/prover/kartlegging-gs/ [The Norwegian directorate of education and training]

Von Aster, M.G., & Shalev, R.S. (2007). Number development and developmental dyscalculia. *Developmental Medicine & Child Neurology, 49*, 868-873.

Vanbinst, K., Ghesquière, P., & De Smedt, B. (2015). Does numerical processing uniquely predict first graders' future development of single-digit arithmetic? *Learning and Individual Differences, 37*, 153-160.

Vanbinst, K., van Bergen, E., Ghesquière, P., & De Smedt, B. (2020). Cross-domain associations of key cognitive correlates of early reading and early arithmetic in 5-year-olds. *Early Childhood Research Quarterly, 52(2)*, 144-152.

VanMarle, K., Chu, F.W., Li, & Geary, D.C. (2014). Acuity of the approximate number system and preschoolers' quantitative development. *Developmental Science, 17(4)*, 492-505.

Van der Ven, S.H.G., Kroesbergen, E.H., Boom, J., & Leseman, P.P.M. (2011). The development of executive functions and early mathematics: A dynamic relationship. *British Journal of Educational Psychology, 82(1)*, 100-119.

Wang, A.H., Firmender, J.M., Power, J.R., & Byrnes, J.P. (2016). Understanding the program effectiveness of early mathematics interventions for prekindergarten and kindergarten environments: A meta-analytic review. *Early Education and Development*, *27*(5), 692-713

Willcutt, E.G., Petrill., S.A., Wu, S., Boada, R., DeFries, J.C., Olson, R.K., & Pennington, B.F. (2013) Comorbidity between reading disability and math disability: concurren psychopathology, functional impairment, and neuropshychological functioning. *Journal of Learning Disabilites, 46(6)* 500-516.

Wilkins, J.L.M., Baroody, A.J., & Tiilikainen, S. (2001). Kindergartner's understanding of additive commutativity within the context of word problems. *Journal of Experimental Child Psychology, 79*(1), 23–26.

Wilson, A.J., Dehaene, S., Dubois, O. & Fayol, M. (2009). Effect of an adaptive game intervention on accessing number sense in low-socioeconomic-status kindergarten children. *Mind, Brain, and Education, 3(4)*, 224-234.

World Health Organization. (2015). ICD-10: Den internasjonale statistiske klassifikasjon av sykdommer og beslektede helseproblemer 2015.

World MedicalAssociation. (2013). WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects. Retrieved 16.11.15, from http://www.wma.net/en/30publications/10policies/b3/index.html

Wright, R. J., Martland, J., Stafford, A. K. (2006). Early numeracy. Assessment for teaching and intervention. 2. Edition. London: Sage.

Zhang, X., Koponen, T., Räsänen, P, Aunola, K., Lerkkanen, M-K, & Nurmi, J-E. (2014). Linguistic and spatial skills predict early arithmetic development via counting sequence knowledge. *Child Development, 85(3)*, pp. 1091-1107.

Östergren, R., & Träff, U. (2013). Early number knowledge and cognitive ability affect early arithmetic ability. *Journal of Experimental Psychology, 115(3)*, 405-421.

Dissertational Studies

I

II

III

I

Study I:

Lopez-Pedersen, A., Lervåg, A., & Melby-Lervåg, M. (Unpublished manuscript). Analyzing the Developmental Relationship between the Approximate Number System and Early Mathematics: Different Methodological Approaches Lead to Different Results.

Study II:

Lopez-Pedersen, A., Mononen, R., Korhonen, J., Aunio, P., & Melby-Lervåg, M. (2020). Validation of an early numeracy screener for first graders. *Scandinavian Journal of Educational Research,* 1-21. https://doi.org/10.1080/00313831.2019.1705901

II

Study III:

Lopez-Pedersen, A., Mononen, R., Aunio, P., Scherer, R., & Melby-Lervåg, M. (submitted). Improving Numeracy Skills in Low Performing First Graders: A Randomized Controlled Trial.

III

Appendix I:

Information letter and request to participation.

I

UiO Institutt for spesialpedagogikk
Det utdanningsvitenskapelige fakultet

Kjære foresatte! Oslo, 21.09.2016

FORESPØRSEL OM DELTAGELSE I FORSKNINGSPROSJEKT

Bakgrunn og formål

Universitetet i Oslo arbeider med et forskningsprosjekt hvor vi ønsker å undersøke hvordan arbeid med matematikk i skolen kan ha positiv virkning på matematikkferdigheter og matematikkutvikling. I den forbindelse skal alle skolene i kommunen delta i prosjektet. Elever på 1. trinn vil få tilbud om å delta. Formålet med prosjektet er tredelt - å vurdere effekten av et intervensjonsprogram i matematikk, lage normer til et screeningsverktøy i matematikk som videre vil kunne bidra til tidlig identifisering av elever som strever i matematikk og å undersøke hvilke faktorer som kan være med på å påvirke utvikling av matematiske ferdigheter.

Hva innebærer deltakelse i studien?

Screening

Elever som deltar skal kartlegges på ulike ferdigheter knyttet til matematikk. Dette gjennomføres i gruppe/klasse av kontaktlærerne, og tar ca. 30-40 minutter. Dette er en skriftlig prøve i matematikk, men er ikke knyttet til elevens vurdering i faget. Formålet med denne screeningen er todelt - både å validere kartleggingsprøven og å bruke den til å velge deltakerne til tiltaket vi ønsker å vurdere effekten av. Etter at screeningen er gjennomført vil omtrent 30 % av elevene bli med videre i prosjektet. Det vil være resultatene på screeningen som avgjør hvem som blir med videre – de elevene som gjennom screeningen viser de trenger tiltaket vi ønsker å vurdere effekten av mest. For å kunne validere screeningen må vi gjennomføre screeningen to ganger i løpet av skoleåret– en i oktober 2016 og en i april/mai 2017, for deretter å sammenligne med kartleggingsprøven fra Utdanningsdirektoratet som alle elevene på 1. trinn gjennomfører. I tillegg vil foreldre motta et spørreskjema om bakgrunnsinformasjon som mattestimulerende tiltak i hjemmet, hjemmespråk og utdanningsbakgrunn

Tiltak

Skolene som disse barna går på vil bli fordelt tilfeldig i to grupper, der den ene halvparten får et pedagogisk opplegg hvor man arbeider med matematikkferdigheter (intervensjonsgruppen) og der den andre gruppen følger vanlig opplæring i skolen (kontrollgruppen). Vi vil trekke lodd om hvem som skal følge tiltaket og hvem som følger vanlig opplegg i skolen.

Lærere ved skolen vil gjennomføre tiltaket, et matematisk intervensjonsprogram, med intervensjonsgruppen tre ganger i uken, to av disse på ca. 45 minutter og en på ca. 10 minutter. Dette varer i 8 uker, etterfulgt av en 6 ukers periode hvor gruppen har en økt per uke som en repetisjonsfase. Det vil tas lydopptak av undervisningen. Intensjonen med dette er å kunne sikre at lærerne følger tiltakets opplegg og instruksjoner, ikke å lytte til elevene. Dersom klassen gjennomgår nytt lærestoff i matematikk som ikke dekkes i intervensjonsmaterialet, vil deltagerne følge klassens undervisning.

Elevene testes før tiltaket starter, etter at 8 uker med tiltak er over, etter at 6 uker med repetisjon er ferdig og ca. et halvt år etter tiltaket, høsten 2017. Dette innebærer utvidet kartlegging innenfor matematikk, språk og generelle evner. Vi ønsker også å følge elevene videre ved å kartlegge dem videre ca. en gang i året fram til slutten av 5. trinn. Dette vil da kun være innenfor matematikk.

Vi ønsker å innhente informasjon om elevens videre skolegang, som for eksempel resultater fra nasjonale kartlegginger. Vi vil sende informasjonsskriv om dette underveis der det minnes om muligheten for å trekke seg om man ikke ønsker øvrig deltakelse. Informasjon om dette kommer fra skolen.

Postadresse: Postboks 1140 Blindern, 0318 Oslo
E-post: ekspedisjonen@isp.uio.no
www.uio.no/isp/

Hva skjer med informasjonen om deg og ditt barn?
All informasjon om ditt barn skal ha koblingsnøkkel, og anonymiseres ved prosjektslutt. Materialet vil bli lagret på eksterne sikrede harddisker og slettet innen oktober 2025. Lydopptakene vil slettes innen januar 2020. Ingen enkeltbarn vil være gjenkjennbare i rapporteringer fra prosjektet. Alle opplysningene behandles konfidensielt og skolen vil ikke få tilgang til testresultatet for hver enkelt elev.

Du har rett til å på forhånd se testmanualer og opplegg som skal forelegges barnet. Du har videre rett til innsyn i hvilke opplysninger som registreres om barnet. Hvis du ønsker innsyn i pedagogisk opplegg, kan du kontakte skolen. Hvis du ønsker innsyn i hvilke opplysninger som registreres, kan du kontakte ansvarlig forsker.

Prosjektet er meldt og godkjent av personvernombudet for forskning, Norsk Senter for forskningsdata AS (NSD).

Frivillig deltakelse
Deltagelse i prosjektet er frivillig. Barn som ikke deltar, vil følge den vanlige skoledagen som normalt. De ansatte på skolen skal være var for elevens signaler og avbryte testingen eller matematikktreningen hvis barnet viser tegn til ikke å ville delta, og dere kan når som helst trekke dere. Dersom dere trekker barnet fra prosjektet, vil alle innsamlede opplysninger bli slettet.

Ta gjerne kontakt med undertegnede på telefonnummer 22 85 44 42/91 83 04 69, anita.lopez-pedersen@isp.uio.no dersom det er spørsmål. Dersom du gir tillatelse til at ditt barn kan delta i prosjektet, er det fint om svarslippen på neste side leveres lærer innen___fredag 30. september 2016_____.

Oppsummert ønsker vi med dette informasjonsskrivet samtykke til følgende:

- Gjennomføring av screening i matematikk, og i arbeidet med å standardisere (validere) denne screeningen gir skolen tillatelse til å utlevere resultatet fra denne for å sammenligne resultatet på de nasjonale kartleggingsprøvene fra Utdanningsdirektoratet

- Deltakelse til tiltaket dersom ditt barn etter screeningen viser seg å kunne ha utbytte av det (dette får dere eksplisitt informasjon om via skolen)

- Mulighet til å følge elevene i tiltaksgruppen ved å innhente informasjon om elevens videre skolegang, som for eksempel resultater fra nasjonale kartlegginger på øvrige trinn

Deltagelse i dette forskningsprosjektet er svært viktig for oss, og vi vil sette stor pris på deres deltagelse.

På vegne av forskergruppen,

vennlig hilsen

__________________________________	____
Anita Lopez-Pedersen	Monica Melby-Lervåg
(Doktorgradsstipendiat ved Institutt for spesialpedagogikk, UiO)	(Professor ved Institutt for spesialpedagogikk, UiO)

SVARSLIPP

Som foresatt har jeg lest informasjonsbrev angående forskningsprosjektet.

Jeg samtykker til at vårt barn

Navn: ____, fødselsdato:________________(dd/mm/åå)

☐ Kan delta i prosjektet.

☐ Kan ikke delta i prosjektet.

Navn på skolen:……………………………………………

Navn og telefonnummer foresatt(e):

____

____

Underskrift____Dato ___________2016

UiO : Institutt for spesialpedagogikk
Det utdanningsvitenskapelige fakultet

Kjære ansatte i skolen!

Oslo, 31.08.2016

FORESPØRSEL OM DELTAGELSE I FORSKNINGSPROSJEKT

Bakgrunn og formål
Universitetet i Oslo arbeider med et forskningsprosjekt hvor vi ønsker å undersøke hvordan arbeid med matematikk i skolen kan ha positiv virkning på matematikkferdigheter og matematikkutvikling. I den forbindelse skal alle skolene i kommunen delta i prosjektet. Barn på 1. trinn vil få tilbud om å delta.

Elever som deltar skal kartlegges på ulike ferdigheter knyttet til matematikk. Dette gjennomføres i gruppe/klasse av kontaktlærerne, og tar ca. 30-40 minutter. Omtrent 30 % av elevene vil bli med videre i prosjektet, det vil være resultatene på screeningen som avgjør hvem som blir med videre – de elevene som gjennom screeningen viser de vil kunne ha utbytte av tiltaket. Skolene som disse barna går på vil bli fordelt tilfeldig i to grupper, der den ene halvparten får et pedagogisk opplegg hvor man arbeider med matematikkferdigheter (intervensjonsgruppen) og der den andre gruppen følger vanlig opplæring i skolen (kontrollgruppen). Elever som ikke deltar i prosjektet følger ordinær undervisning, uavhengig hvilken klasse de går i.

Hva innebærer deltakelse i studien for ansatte på skolen?
Gjennom loddtrekning vil klassene som de aktuelle barna går i fordeles tilfeldig i en tiltaksgruppe og en kontrollgruppe. Kontrollgruppen følger ordinær undervisning i skolen. Pedagogisk personale på skolen som blir med i tiltaksgruppen vil gjennomføre tiltaksprogrammet. Skolens ledelse og lærere har deltatt på informasjonsmøte om studien, og den enkelte skole avgjør hvem som skal gjennomføre tiltaket. Hvis du ønsker å delta i forskningsprosjektet, ved å gjennomføre det pedagogiske opplegget vårt, leverer du samtykkeskjema til rektor som formidler det videre til forskerne. Ansvarlig(e) pedagog(er) vil få opplæring av forskerne i prosjektet. Dette vil foregå som to halvdagsseminarer i perioden november-desember 2016.

Prosjektet starter oktober/november 2016, og tiltaket i skolen vil pågå fra og med januar 2017 og til og med våren 2017. Lærere ved skolen vil gjennomføre et matematisk intervensjonsprogram med intervensjonsgruppen tre ganger i uken tiltaksperioden (to av disse på ca. 45 minutter og en på ca. 10 minutter). Det vil tas lydopptak av undervisningen. Intensjon med lydopptak er å sikre gjengivelse av studiens og undervisningsøktenes instruksjoner. Dette er informasjon forskerne trenger for å kunne vurdere i hvilken grad opplegget har blitt fulgt når man skal vurdere resultatene når studien er avsluttet, ikke for å gi veiledning, tilbakemelding eller som vurdering av lærerne.

Hva skjer med informasjonen som innsamles?
I tillegg til lydfilene ønsker vi å samle informasjon om lærernes utdanningsbakgrunn og kjenn, både for elevene i tiltaksgruppen og elevene i kontrollgruppen. Vi ønsker videre at lærerne gjennomfører enkel loggføring underveis i prosjektet. Dette vil bli lagret på eksterne sikrede harddisker og slettet innen januar 2025. Alle opplysningene behandles konfidensielt.

Du har rett til innsyn i hvilke opplysninger som registreres om deg som ansatt. Hvis du ønsker innsyn i hvilke opplysninger som registreres, kan du kontakte ansvarlig forsker.

Prosjektet er meldt til personvernombudet for forskning, Norsk Senter for forskningsdata AS (NSD).

Frivillig deltakelse
Deltakelse i prosjektet er frivillig. Ansatte som ikke ønsker å delta, vil følge den vanlige skoledagen som normalt. Både forskere og ansatte i skolen skal være var for elevenes signaler og avbryte testing eller matematikktrening hvis barnet viser tegn til å ikke ville delta.

Dere kan når som helst trekke dere fra prosjektet uten at dette får konsekvenser. Dersom dere trekker dere fra prosjektet, vil alle innsamlede opplysninger bli slettet.

Ta gjerne kontakt med undertegnede på telefonnummer 22 85 44 42/91 83 04 69 dersom det er spørsmål.

På vegne av forskergruppen,

vennlig hilsen

________________________________ ________________________________

Anita Lopez-Pedersen
(Doktorgradsstipendiat ved Institutt for spesialpedagogikk, UiO)

Monica Melby-Lervåg
(Professor ved Institutt for spesialpedagogikk, UiO)

Som ansatt har jeg lest informasjonsbrev angående forskningsprosjektet.

Jeg samtykker til å delta i prosjektet

Underskrift____Dato _______________

Appendix II:

Ethical approval from the Norwegian Centre for Research Data (NSD)

II

Anita Lopez-Pedersen
Institutt for spesialpedagogikk Universitetet i Oslo
Postboks 1140 Blindern
0318 OSLO

Vår dato: 19.09.2016 Vår ref: 49265 / 3 / AGH Deres dato: Deres ref:

TILBAKEMELDING PÅ MELDING OM BEHANDLING AV PERSONOPPLYSNINGER

Vi viser til melding om behandling av personopplysninger, mottatt 15.07.2016. Meldingen gjelder prosjektet:

49265	*The effect of mathematical intervention programs for children performing low in mathematics.*
Behandlingsansvarlig	*Universitetet i Oslo, ved institusjonens øverste leder*
Daglig ansvarlig	*Anita Lopez-Pedersen*

Personvernombudet har vurdert prosjektet og finner at behandlingen av personopplysninger er meldepliktig i henhold til personopplysningsloven § 31. Behandlingen tilfredsstiller kravene i personopplysningsloven.

Personvernombudets vurdering forutsetter at prosjektet gjennomføres i tråd med opplysningene gitt i meldeskjemaet, korrespondanse med ombudet, ombudets kommentarer samt personopplysningsloven og helseregisterloven med forskrifter. Behandlingen av personopplysninger kan settes i gang.

Det gjøres oppmerksom på at det skal gis ny melding dersom behandlingen endres i forhold til de opplysninger som ligger til grunn for personvernombudets vurdering. Endringsmeldinger gis via et eget skjema, http://www.nsd.uib.no/personvern/meldeplikt/skjema.html. Det skal også gis melding etter tre år dersom prosjektet fortsatt pågår. Meldinger skal skje skriftlig til ombudet.

Personvernombudet har lagt ut opplysninger om prosjektet i en offentlig database, http://pvo.nsd.no/prosjekt.

Personvernombudet vil ved prosjektets avslutning, 30.10.2025, rette en henvendelse angående status for behandlingen av personopplysninger.

Vennlig hilsen

Kjersti Haugstvedt

Agnete Hessevik

Kontaktperson: Agnete Hessevik tlf: 55 58 27 97
Vedlegg: Prosjektvurdering

Dokumentet er elektronisk produsert og godkjent ved NSDs rutiner for elektronisk godkjenning.

Errata list

Name of candidate: Anita Lopez-Pedersen

Title of thesis: On the Trail of Early Numeracy Skills. Understanding, identifying and ameliorating young children's early numeracy skills: A multimethod approach

Abbreviations for different types of corrections:

Cor – correction of language

Cpltf – change of page layout or text format

Page	Line	Original text	Type of correction	Corrected text
4	19	30 – 40 %	Cpltf	30-40%
6	21	in calculation (Dehaene &	Cpltf	in calculation (Dehaene &
11	6	. Additionaly, this	Cpltf	. Additionaly, this
12	2	Rittle-Johnson	Cpltf	Rittle-Johnson
14	9	Hoard (2005)	Cpltf	Hoard (2005)
22	22	Resnick 1989	Cpltf	Resnick, 1989
32	6,7, 14	Pre-existing	Cor	Preexisting
40	31	the Early Numeracy screener	Cor	the Early Numeracy Screener
41	23	The Early Numeracy screener	Cor	the Early Numeracy Screener
42	27	the Early Numeracy screener	Cor	the Early Numeracy Screener
12 (dissertation paper 1)	22	Lopez-Pedersen, Korhonen, Aunio, Mononen, & Melby-Lervåg, 2020	Cor	Lopez-Pedersen, Mononen, Korhonen, Aunio, & Melby-Lervåg, 2020
18 (dissertation paper 1)	9	at the between level	Cor	at the within level
32 (dissertation paper 1)	10	âˆ'.07	Cor	.07