Early Numeracy is the Important Factor for the Development of Arithmetic, not Number Sense

A longitudinal study about the role of number sense in arithmetic

Ingvild Elise Selfors Reppe

Master thesis at the Department of Special Needs Education
Faculty of Educational Sciences

UNIVERSITETET I OSLO

Spring 2019

Early Numeracy is the Important Factor for the Development of Arithmetic, not Number Sense

A longitudinal study about the role of number sense in arithmetic at age.

2019

Early Numeracy is the Important Factor for the Development of Arithmetic, not Number Sense

Ingvild Elise Selfors Reppe

http://www.duo.uio.no/

Trykk: Reprosentralen, Universitetet i Oslo

Summary

Mastery of basic arithmetic skills is a key goal of early education. At this point, the understanding of the underlying cognitive factors for the development of arithmetic abilities is restricted. The literature reports different findings on the importance of the cognitive factors. Some studies found that the preverbal ability number sense has a strong relation with arithmetic (Bonny & Lourenco, 2013; Desoete et al., 2012; Libertus et al., 2011; Libertus et a., 2013; Mazzocco et al., 2011; Toll et al., 2015) whereas other studies report that the important cognitive factors are the abilities within early numeracy (Göbel et al., 2014; Sasanguie, et al., 2014). Because of the conflicting findings in the literature, this study aimed to investigate to what extent number sense in 5-year-olds can predict arithmetic skills one year later in a sample of Norwegian children. To answer this the research questions are as followed:

- To what extent can number sense in children at age five predict arithmetic ability halfway through first grade?
- Is there still a relation when controlling for early numeracy, age, vocabulary and non-verbal intelligence?

The study is part of the longitudinal research project *Development of Numeracy and Literacy in Children (NumLit)* at the Department of Special Needs Education of the University of Oslo. This study has a quantitative approach within a non-experimental observational longitudinal design. The sample consists of 187 children, participating at age 5 in Kindergarten and again at age 6 in 1st grade.

For measuring number sense a dot comparison test from TOBANS (Brigestocke et al., 2016) was used. Early numeracy was measured with the tests 'Number reading' and 'Count on', both specifically made for the NumLit project. Arithmetic was measured with an addition fluency and a subtraction fluency test from TOBANS (Brigestocke et al., 2016). The covariate non-verbal intelligence was measured with *The Raven Colored Progressive Matrices* (Raven, 1998) and the covariate vocabulary was measured with *British Picture Vocabulary Scale* (BPVS II).

The data was analyzed using SPSS for descriptive, correlational and hierarchical multiple regression analysis. The analysis were conducted to assess if there were a relation

between the dependent and independent variables and to what extend the predictor variables could explain unique variation in arithmetic.

The results of the analysis in this study showed that number sense at age 5 in Kindergarten has a weak, but significant correlation with arithmetic one year later. The same holds for the control variables non-verbal intelligence and age. Importantly, there was a moderate significant relation between early numeracy and arithmetic. The regression analysis indicated, that the relation between number sense and arithmetic was not predictive in nature, while controlling for non-verbal intelligence and age. Only early numeracy had a significant on arithmetic and explained unique variance. Therefore, it can be concluded that the important underlying cognitive factor at age 5 for predicting arithmetic is early numeracy.

Preface

First, a great thanks to my supervisor Sietske van Viersen for good guidance and constructive feedback during the writing process. Thank you for always being available for questions, without your guidance and patience, this thesis would not have been the same. I also want to thank the research team *Development of Numeracy and Literacy in Children (NumLit)* at the Department of Special Needs Education at the University of Oslo for giving me the opportunity to participate in their research program and for allowing me to use their collected data for this thesis.

Further, I want to thank my mom for always being available and for the good discussion during this semester and asking critical questions. A great thanks to my brother, Jan-Kåre for proofreading, and not least to my dad for good hot meals and encouraging words.

Last but not least, I want to thank my fellow students for the cooperation and good coffee breaks.

June 2019,

Ingvild Elise Selfors Reppe

Content

Figure overview

Table overview

1 Introduction

To function in modern society, we are required to process exceptional levels of numerical information. We need to handle information from computers, smartphones, about finances and healthcare, which all require a form of numerical fluency (Prince & Ansari, 2013). In addition, we use numeracy when shopping, while avoiding the longest queue in the supermarket and by making a quick judgment without counting (Bonny & Lourenco, 2013). Of course, when we use money, or while baking, we rely on our math skills. The mathematical abilities pupils have when they start school have proven to be of great importance for mathematical development throughout their education. These abilities are essential for future career success and for becoming an active contributing participant in society (Geary, 2015; Price & Ansari, 2013). Preverbal math abilities and early numeracy have proven to be persistent predictors of mathematical achievement (Aunola, Leskinen, Lerkkanen, & Numri, 2004; Libertus, Feigenson, & Halberda, 2011). For instance, numerical knowledge at age 7 predicts socioeconomic status at age 47, even after controlling for a wide range of relevant variables (Ritchie & Bates, 2013).

Today's pupils are our future citizens, and every pupil should be given equal opportunity to contribute to society in the future (Kunnskapsdepartementet, 2006, p. 11-12). Having a basic understanding of mathematical concepts and numbers is a prerequisite for this. In Norway calculation is one of the five basic skills, along with reading, writing, IT and oral skills. These skills are included in all school subjects in Norway (Udaninngsdirektoratet, 2017, p. 2). This means that if you struggle with math you might struggle in other subjects as well. According to OECD (2014), one must score at level 2 or higher on the PISA test to become a successful citizen. PISA results from 2015 show that 17 % of Norwegian 10th graders scored at level 1 (OECD, 2014; OECD, 2016). This means that almost 20 % of the students that year are predicted not to graduate high school and are at a high risk of falling outside the community (Ducan et al., 2007). In recent years, there has been a greater focus on increasing the competence of teachers, with a focus on early intervention. According to the results from TIMSS 2015, this may have yielded good results. Norwegian 4th and 5th graders scored highest when compared to their peers in the Nordic countries and were among the best in Europe (Bergem, Kaarstein, & Nilsen, 2016). Knowing that numerical competence in early life is a strong predictor of later achievements in life, a big responsibility falls on the school and teachers, to discover children that are struggling or are at risk of developing mathematical difficulties. Therefore, it is important that teachers can recognize students at risk for developing mathematical difficulties at an early age, to prevent students from falling behind and in the worst case, dropping out of school (Korhonen, Linnanmäki, & Aunio, 2014).

The reasons why some pupils struggle with mathematics vary and there are major individual differences. There are different factors that can play a critical role in learning mathematics. These can be cognitive factors, such as non-verbal intelligence, working memory, executive functions (Arán, Filippetti, & Richaud, 2017; Andersson & Lyxell, 2007; Stock, Desoete, & Royers, 2009), and/or deficits in processing numerical magnitude information (Geary, 2000). In addition, it can be a result of external factors, such as poor teaching, low socio-economic status or behavioural and/or attention difficulties (Butterworth, Varma, & Laurillard, 2011; Geary, Hoard, Nugent, & Byrd-Craven, 2008; Price & Ansari, 2013). In this study the focus is on cognitive factors. Among the cognitive factors there is one that is highly debated in the literature and possibly crucial in learning mathematics; number sense (NS; Libertus et al., 2011; Halberda & Feigenson, 2008). Number sense is the inherent, quick and intuitive ability to compare amounts or quantities without counting (Aunio & Räsänen, 2015). This skill is thought to be fundamental in acquiring mathematics (Brannon & Park, 2015; Starkey, 1992). Thus, number sense lays the foundation for learning how to count, which is necessary for developing an understanding of basic arithmetic. Mathematics as a school subject has a cumulative structure, which means that new skills are built on more fundamental abilities (Aunola et al., 2004; Jordan, Kaplan, Ramineni, & Locuniak, 2009). Therefore, pupils who do not master one element, can experience difficulties in acquiring new elements, which in turn can result in academic challenges later on (Hornigold, 2015). Today, there are uncertainties as to which abilities are critical in the acquiring of arithmetic abilities. Hence, this master thesis aims to look further into premath abilities and their predictive role for basic arithmetic in school.

1.1 Purpose

The aim of this study is to investigate the significance of number sense for the development of arithmetic skills. With data obtained from monolingual 5-year-olds Norwegian children in Kindergarten, this study will, investigate whether number sense can predict arithmetic abilities in 1st grade. Since this study is longitudinal, with data at two time points, the study can examine change and development within the sample over time. To ensure that a relationship between number sense and arithmetic is not influenced by other factors affecting mathematical development, early numeracy, age, non-verbal intelligence and vocabulary will be controlled for.

- To what extent can number sense in children at age five predict arithmetic ability halfway through first grade?
- Is there still a relation when controlling for early numeracy, age, vocabulary and non-verbal intelligence?

Based on the conflicting findings in current research and lack of studies conducted on Norwegian children, this study will try to add clarifying information about ANS role for arithmetic abilities in Norwegian children. This study will therefore be a contribution to the research of which underlying cognitive factors plays a predictive role for arithmetic abilities.

1.2 Limitation and terms

This master thesis is associated with a larger longitudinal research project that has given access to a good test-battery and a large sample size. The sample consists of monolingual Norwegian-speaking children that were followed over the time span of one year. The thesis contains the following variables; non-symbolic comparisons skills, counting skills, number reading, addition fluency, subtraction fluency, non-verbal intelligence and receptive vocabulary. The variables selected are chosen to measure number sense, early numeracy, early arithmetic, intelligence and vocabulary. This choice of concepts is made on the theoretical and empirical framework in which the study is placed, to investigate the role of number sense in the development of arithmetic, while also taking early numeracy, intelligence, vocabulary and age into account.

Important terms are defined the first time they are used, and some are abbreviated after first use. It is assumed that the reader has some degree of prerequisite knowledge within the pedagogical and/or special needs education field, as it is not possible to go in depth for all topics. The thesis sometimes refers to the participants as children, and sometimes as pupils. This is because of the two-time point- in the first, the participants were in kindergarten and are by definition children, while in the second time point, they are pupils at school.

1.3 The thesis structure

The introduction contains some of the thesis background and the main research questions. Part 2 covers the theoretical framework of the thesis. It starts with a review of the development of mathematical abilities. Subsequently it takes a look at mathematical difficulties and subtypes, before going into the triple code model. From here we take a closer look at number sense, and approximate number sense, and the role of these skills in the development of arithmetic abilities as based on other studies in the literature. Part 3 covers the methodological approach of the study, the study's design and measures and ethical considerations. Part 4 presents the analysis, and the results are described. In part 5, the study results are discussed in light of the theory and empirical findings in order to answer the thesis hypothesis and the research questions. The validity and reliability of the study are considered, as are the study's implications for education and the way forward.

2 Theoretical and empirical background

Today there are many different views on how arithmetic abilities develops and why some children develop difficulties within mathematics (Dehaene, 1992; Geary, 2000; Göbel, Watson, Lervåg, & Humle, 2014; Mazzocco, Feigenson, & Halberda, 2011). Studies find different predictors for mathematical development, such as executive function, intelligence, early numeracy and number sense. This study aims to investigate to what extent some of these skills can predict arithmetic abilities. Therefore, theories on how mathematical abilities and deficits develops are included. This chapter explains fundamental mathematical abilities and their development, by focusing on the early development. To illustrate this, mathematical difficulties and subtypes of dyscalculia are also addressed before we go into the innate ability that has been hypothesized to form the basis, i.e. number sense, and its role for arithmetic.

2.1 Mathematical development

There are different factors that are a prerequisite for learning and mastering mathematics. In general, these factors can be divided into domain-general and domain-specific precursors. The domain-general precursors consist of general cognitive abilities, such as working memory, processing speed, and intelligence. These factors have been found to be important when acquiring mathematics, and also in other school subjects (Passolunghi & Lanfranchi, 2012). For example, intelligence has, among others, proven to have a direct relation with mathematical abilities. Since these are all factors that affect more skills than just mathematical abilities, they are considered general factors. Domain-specific precursors include abilities that are only required for or mainly affect mathematical acquisition, for example, number sense and early numeracy (Passolunghi & Lanfranchi, 2012). This study mainly focuses on domain-specific precursors, that can predict arithmetic abilities. Geary (2000) divides the domain-specific abilities up and distinguishes between primary and secondary skills. Primary abilities are the biological quantitative abilities that are innate, as an intuitive perception, and are involved in the preverbal number system. In this system, abilities are found as early as infancy and they provide a skeletal structure that is necessary for more complex numbers, counting and arithmetic skills in preschool (Geary, 2000). The preverbal number system develops along with the child (Geary, 2000; Starkey, 1992|) and devlopment of these abilities is considered to follow a universal normative pattern (Geary, 2000). Secondary abilities are known school-taught competencies, that are culturally, and generation bound. A result of this, is that there is no normative pattern for the development of secondary skills (Geary, 2000). For example, most children in Norway go to kindergarten from the year they one until the year, they turn six and start school. The Norwegian framework plan for mathematics in kindergarten (Kunnskapsdepartementet, 2017, p. 53-54), states that the kindergarten phase is about discovering, exploring and creativity, to help the child understand mathematics. The kindergarten is supposed to help the child discover and wonder about mathematical relationship, in a way that helps the child to develop a fundamental understanding of the mathematical concepts, and, to play and experiment with numbers, quantities and counting. Kindergarten staff should help the children accomplish this through books, games, music, digital tools, toys and materials (Kunnskapsdepartementet, 2017, p. 53-54). The Norwegian framework is fairly open for interpretation and each kindergarten makes their own plan on how to achieve these goals. This can result in children in Norway acquiring mathematical abilities to differing degrees before reaching school age. In other countries, there can be a more, or less concrete framework regarding the acquisition of early mathematics, even though there are some similar developmental patterns across nations and generations (Geary, 2000). Based on Geary (2000), some early skills that develop during preschool lay the foundation for learning more complex mathematics. In the chosen theory of numerical development, it is assumed that the understanding of numerical magnitudes increases with age, and that one ability builds upon the previous. Below, we will look at what abilities these are and how they are assumed to develop. Here, they are divided into preverbal abilities, early numeracy and arithmetic.

2.1.1 Preverbal abilities

Preverbal abilities are abilities that do not require language knowledge, that infants develop before they learn to talk. Preverbal abilities are basic quantitative abilities that are innate, which can be found at different stages of infancy, and later on during development. These are abilities such as understanding numerosity, ordinality, and simple arithmetic (Geary, 2000). Numerosity is the ability to accurately determine the quantity of a small set of items, without counting. Ranging from 1 to 4 items, this is also known as subitizing, the intuitive and precise recognition of a quantity of objects. Beyond that, it is the ability to approximately assess quantities without counting, determining which quantity is most numerous (Aunio &

Räsänen, 2015; Libertus et al., 2011). For example, infants can determine amounts with a 1:2 ratio (Halberda & Feigenson, 2008) which is often illustrated with dot-comparison tasks. Ordinality is the understanding of the concepts 'more than' and 'less than'. Early one, this is limited to quantities that are not greater than 5, infants know for example that 3 is more than 2, but less than 4. When they are about 5 months old, infants are already sensitive to changes of quantity in small sets. They react e.g. with longer gaze duration, when items are added or subtracted from a set of 2 items (Geary, 2000). This can be considered as early arithmetic understanding.

These abilities are believed to provide a foundation that are necessary for acquiring more complex number knowledge. This means that the preverbal number system is integrated into the child's increasing language competencies. For example, an infant of approximately 18 months can count up to 3 or 4 physical objects with number words (Geary, 2000). That take us to early numeracy.

2.1.2 Early numeracy

Early numeracy is the growing ability to understand and make relational statements about mathematics. These are abilities such as counting consisting of comparing, classification, seriation and one-to-one correspondence, for understanding ordinality and cardinality (Aunio & Niemivirta, 2010; Desoete, 2015). The development for early numeracy begins with the acquisition of whole-number-word sequences and this can be divided into 6 stages. It starts with a primary understanding, that number words correspond to a number of objects. Next, they learn that number words are in a sequence, but they cannot count in the correct order. The subsequent stage is when they can say the number-words in the correct order and at the same time point to objects. About 6 months later they are able to do this with one-to-one accurate correspondence. At the last stage, they can say number words correctly, starting at one and count with one-to-one correspondence, while they understand that the last number said, indicates the total number of objects, i.e., cardinality (Aunio & Niemivirta, 2010). As such, counting abilities can be divided into three aspects; knowledge of number symbols, numerical order and enumeration, each of which are necessary for developing counting strategies that can be used to solve new and unknown mathematical tasks, such as adding up and counting on (Anuio & Räsänen, 2015).

Early numeracy also includes knowledge of Arabic numerals, such as combining the number word 'four' with the Arabic numeral '4', which in turn corresponds with four dots or items (Siegler & Braithwaite, 2004). This includes abilities such as number identification, number reading and symbolic number comparison. Primary school children are expected to master the counting system, get an understanding of the base-10 system, and learn to transcode numbers from their different representations (i.e., number words to Arabic numerals) upon school entry in almost every industrialized nation (Geary, 2000).

2.1.3 Arithmetic

Arithmetic abilities start to develop early and result from counting strategies like adding 1 to 3, where the child puts up 3 fingers plus 1 finger and then counts all of them. It starts with counting-based strategies and after a counting strategy has been successful multiple times, the outcome becomes automatized, and can then be retrieved from memory without having to apply the counting strategy, i.e., arithmetic fact retrieval. When children use other strategies than fact retrieval, they choose the fastest back up strategy for the specific calculation (Siegler & Braithwaite, 2004). The first arithmetic skills that are mastered are addition and subtraction, which lay the foundation for further mastery of multiplication and division (Aunio & Räsänen, 2015). Arithmetic skills are acquired through verbal tasks, text tasks and problem-solving tasks. The ability to conduct arithmetic tasks becomes faster and more accurate with exposure, ultimately resulting in automated knowledge of arithmetic facts.

Like counting, solving arithmetic tasks in word problems are a part of the knowledge children should gain during primary school. The extent that children learn basic symbolic arithmetic facts and computational procedures for solving complex arithmetic problems, varies. With sufficient practice, almost every typically developing child can learn to memorize basic arithmetic facts. However, lack of training can result in retrieval errors and prolonged use of counting strategies (Geary, 2000).

2.1.4 Summary

Early mathematical abilities develop one after another and start with innate intuitive numerosity perception, ordinality, counting, and simple arithmetic. These skills lay the foundation for learning number words, ordinal relationships, and cardinality, which constitute early numeracy. Early numeracy in turn forms the foundation of more complex arithmetic tasks and understanding mathematical symbols, place value and the decimal number system.

Some studies have shown that there is a significant relationship between children’s early numeracy and more complex arithmetic abilities (Aunio & Niemivirita, 2010; Göbel et al, 2014; Jordan, Kaplan, Nabors, & Locuniak, 2006), while others have found a relationship between number sense and arithmetic abilities (Libertus et al., 2011; Mazzocco et al., 2011; Geary, 2013; Toll, van Viersen, Kroesbergen, & van Luit, 2015) Children with mathematical problems lag behind their peers in both of these early math concepts, and the reasons why mathematical problems arise are numerous. To gain more insight into the roles of number sense and early numbers, mathematical difficulties are accounted for.

2.2 Mathematical difficulties

Sometimes math development is atypical, and difficulties may arise. Previous studies have shown that the societal cost of poorly developed mathematical skills is higher than that of poor reading skills, and that those with mathematical difficulties have a poorer future prospect (Geary, 2015). Interest in mathematical difficulties has increased nationally and internationally in recent years. It is a relatively young field that is characterized by various explanations and definitions in the academic literature and in the field (Price & Ansari, 2013). A product of the burgeoning research in this area is the distinction between domain-general deficits and domain-specific deficits, underlying mathematical difficulties (Passolunghi & Lanfranchi, 2012).

Domain-general deficits emphasize cognitive abilities such as working memory, processing speed and intelligence as the main source of the development of math difficulties (Passolunghi & Lanfranchi, 2012). Domain-general deficits lie within the so called secondary abilities. Secondary refers to mathematical deficits stemming from factors outside the math domain, such as poor teaching, low socio-economic status, behavioral, attention and domain-general cognitive deficits (Price & Ansari, 2013). In contrast domain-specific deficits emphasize a lack in understanding of numerosity, ordinality, counting and simple arithmetic. This concerns the primary biological-based quantity ability and is related to impaired development of brain mechanisms for processing numerical magnitude information (Geary, 2000; Passolunghi & Lanfranchi, 2012; Price & Ansari, 2013). Whether the deficit stems from domain-general or domain-specific causes, children with mathematical difficulties often have problems with establishing and automatically retrieving arithmetic facts, lack of numerical facility, and knowing basic number facts by heart (Desoete, Stock, Schepens, Baeyens & Royers, 2009). When mathematical problems become severe and specific, they can be considered at a clinical level and diagnostic assessment is necessary.

2.2.1 Diagnosis

Children that are on the bottom quartile of a standardized mathematics test should be considered low-achieving (LA) and at risk of developmental dyscalculia (DD). LA children have a mild but persistent learning difficulty in mathematics and scores between the 25^{th} and the 11^{th} percentile on standardized mathematical tests for at least two consecutive academic years. Those that score at or below the 10^{th} percentile can be diagnosed with DD (Geary, 2015).

The diagnostic system DSM-V defines developmental dyscalculia as a Specific Learning Disorder characterized by impairments in mathematics (American Psyhiatric Association, 2013). The difficulties should persist for at least 6 consecutive months and should be within the area of numeracy, learning numerical facts, ability to accurately and fluently calculate and accurately reasoning (American Psyhiatric Association, 2013). In Norway ICD-10 is used and denotes the issues above as a lack of the ability to master basic operations; addition, subtraction, multiplication and division (World Health Organization, 2015). Both diagnostic systems use a definition where the abilities must be considerably weaker than what's expected for the child's age, and this should not result from other difficulties, impairments or a lack of teaching (American Psyhiatric Association, 2013; World Health Organization, 2015).

2.2.2 Occurrence

The prevalence of mathematical difficulties depends on the definition applied. Numbers vary from 4 – 14 % for developmental dyscalculia and 15 – 25 % for generally low-achievement (Geary D. , 2015b). Variation in these figures may come from unclear use of terms and definitions and that difficulties are often categorized by severity, ranging from low achievement to mathematical learning disabilities (MLD) and developmental dyscalculia. However, there are no clear cut-off points that define these groups or guide the use of these definitions. Still, there is a consensus that mathematical difficulties are about a serious failure in the ability to acquire arithmetic abilities (Butterworth et al., 2011).

2.2.3 Developmental Dyscalculia and subtypes

Developmental dyscalculia (DD) is a specific learning disability that severely affects the acquisition of arithmetic skills. DD appears to be, like other learning disabilities, a brain-based disorder, that is characterized by a deficit in basic numerical skills (Von Aster & Shalev, 2007). Deficit in areas such as counting, number transcoding, magnitude comparison, mental calculation, placing Arabic numbers on an analogue number line, perceptual quantity estimation and contextual estimation have resulted in three different subtypes of DD; pervasive subtype, verbal subtype and the Arabic subtype (Von Aster, 2000).

The verbal subtype presents a disorder in linguistic processing. It can cause difficulties with the verbal representation of numbers such as counting, the use of linguistically based arithmetic processes, such as mental calculation and retrieval of arithmetic facts. The Arabic subtype represents difficulties with the acquisition of Arabic numerals. This includes difficulties with reading, writing and comparing Arabic numerals. This can appear as having difficulties using the space value system. The pervasive subtype represents a lack of understanding of the primary concept of numbers and numerosity. This implies an inability to develop number representations as a mental number line (Von Aster, 2000). These subtypes are based on the theoretical framework developed by Dehaene (1992) and will be described in more detail below.

2.3 The triple code model

In 1992, Dehaene proposed the triple code model (TCM) for numerical processing, consisting of three modules in which numbers can be represented and manipulated. Numerical processing is a complex ability that consist of mechanisms such as understanding arithmetic principles or memorizing and retrieving arithmetic facts. It emerged as a clear-cut distinction between the system that does the approximate calculation of quantities and a second calculation system based on memorized facts. Based on this the TCM is a multirooted model of numerical processing. The model proposes that there are three codes that serve different functions, and that every numerical procedure is tied to a specific input and output code. These functions have distinct neural architecture and are related to specific tasks (Dehaene & Cohen, 1995). Here, the triple code model is presented as a diagram showing the processes underlying arithmetic expertise and their interaction (see Figure 1).

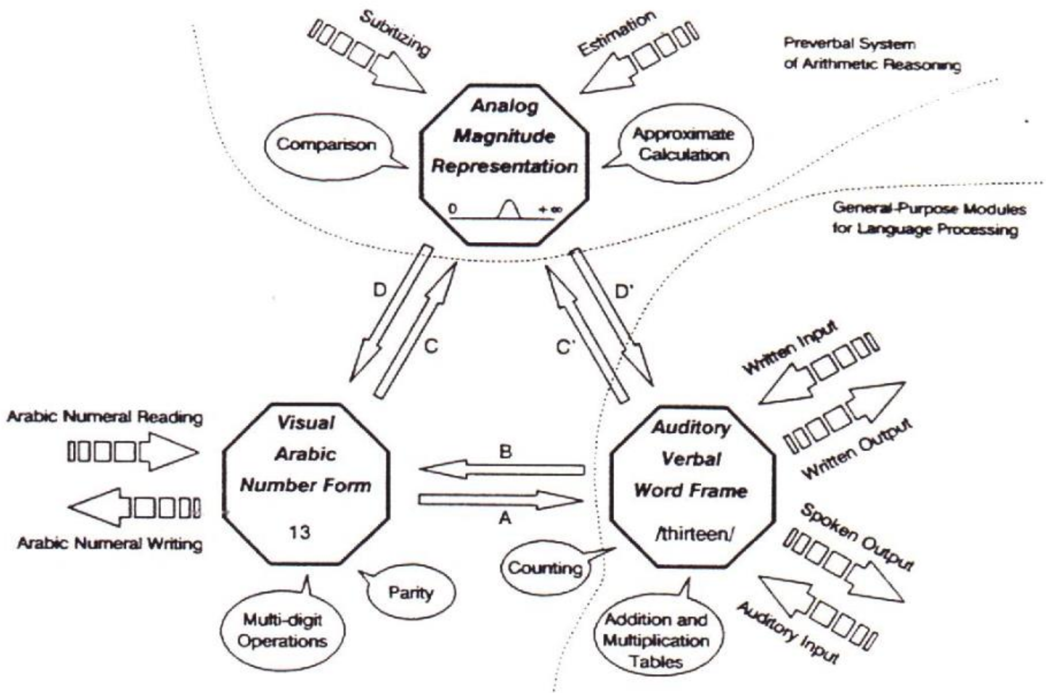

Figure 1 The triple code model (Dehaene, 1995).

The analogue magnitude code contains abilities such as approximating numerosities, number comparison, and placing numerical quantities on a mental analogue number line (Dehaene, 1992; Von Aster, 2000). The preverbal analogue magnitude code develops first (Toll & Van Luit, 2013). This suggests that children are born with an innate numerical ability that is needed to understand how and why number-words, basic numerical principles, counting principles and arithmetical concepts are used. In the absence of these inherent semantic roots, the different concepts are learned without a basis, and are therefore considered to be part of DD when deficient, caused by genetic factors influencing brain development or early brain damage (Von Aster 2000).

The auditory verbal code includes abilities such as counting, use of counting in addition and subtraction, and fact retrieval (Von Aster, 2000). The auditory verbal code is considered to be a part of general language and is also used for the alphabet and knowledge of, for example, the months. The auditory verbal code starts developing in early childhood (Toll & Van Luit, 2013). Children can often count as if they are singing a song without comprehending the one-to-one correspondence.

The visual Arabic code incudes abilities to manipulate and represent numerical quantities in Arabic format. This includes reading, writing and comparing Arabic numerals (Dehaene, 1992). Verbal and visual experience with Arabic numerals lays the foundation for early math competence, and is gradually integrated with the innate non-verbal knowledge, resulting in complex cognitive representations, number words are connected to quantity representations (Toll & van Luit, 2013).

The modules in the triple-code model seem to emerge and be elaborated upon at different times during development. The first stage concern inborn universal numerical abilities, a genetic predisposition (Geary, 2000; Von Aster, 2000). Followed by a sematic analogue module that is necessary for understanding basic counting, numerical principles and the arithmetic concept; the verbal module. The verbal module in turn lays the foundation to learn the connection between verbal numbers and Arabic numerals, and to be able to read, write and compare Arabic numerals. These three modules keep developing and are the foundation for the next steps of comprehending and mastering mathematics (Dehaene, 1992; Geary, 2000; Von Aster, 2000). For access to the different modules there are two main routes. One that transcodes written numerals to auditory verbal representations, which is used for retrieval of arithmetic facts without semantic mediation. The other is an indirect sematic route, that is specialized in quantitative processing by manipulating analogue magnitude representations. This route compares numerical (operands) and uses back-up strategies by manipulating visual Arabic representations, when verbal memory knowledge is not available, i.e., 14 + 5 = 10 + 4 + 5.

Schmithorst and Brown (2004) designed a functional magnetic resonance imaging (fMRI) paradigm involving mental addition and subtraction of fractions, to explore the triple-code model. They tested educated adults with stimuli that involved all three modules. They found that the TCM of analogue magnitude, auditory verbal and visual Arabic code is a good framework for analyzing the performance of complex mathematical tasks. Based on the model, it can give a better understanding of genetic disorders in number processing, like DD and subtypes of mathematical disabilities. Many studies have focused on the most basic module of the triumvirate, the analogue magnitude code, which we'll have a further look at in the next part.

2.4 Number sense

Numbers are a fundamental parameter of elementary abilities for making sense of the world (Dehaene, 2001). This ability is the skill to mentally represent and manipulate numerosities on a number line, which is an analogue representation of numbers. In short, it is the innate ability to quickly understand, approximate, and manipulate numerical quantities (Dehaene, 2001, p. 16-17). This is known as the number sense (NS) and is a part of the domain-specific precursors (Dehaene, 2001), the biological primary abilities (Geary, 2000) and the analogue magnitude code (Dehaene, 1992). There is evidence that supports the existence of a biologically inherited number sense; its presence in animals and human infants and the homology between them, and that it has a neural substrate in a specific area of the cerebellum. This will be briefly reviewed below.

The NS is believed to be a biologically inherited ability, which is partly under genetic control (Dehaene, 2001). Based on this, there should be a precursor ability in animals as well. Therefore, several studies have searched for evidence of this, and found that among others, mammals, birds, amphibians and fish, are able to discriminate between quantities. For example, lionesses decide whether or not to approach other lionesses based on how many lionesses roars they hear. Similarly, fish prefer to join large shoals when meeting predators. This indicates that animals are able to discriminate the larger quantity of conspecifics (Agrillo, 2015). Animals have been shown to discriminate between quantity of sets, both by visual objects and auditory sequences of sounds. Other studies have also found that animals have internal representations of numbers, and are able to perform approximate addition, subtraction and comparison tasks (Dehaene, 2001). These findings indicate that the NS is a part of the biological primary quantitative abilities.

NS found in animal is highly identical to that found in preverbal infants. Studies have found that 6-to-7-month-old infants are able to discriminate between visual quantities (Dehaene, 2001), that preverbal infants can discern small and large numbers of objects, as well as apply ordinality, enumeration, counting and simple arithmetic (Geary, 2000; Mccrink & Birdsall, 2015). This can be divided in to two systems. One is the object-file system, were they can inutility and precisely recognize small amounts of 1-4 objects. The large-number system represents the ability to approximate the numerical magnitude of large quantities (Mccring & Bridsall, 2015; Starkey, 1992).

From the two previous sections there are two characteristics that humans and animals share; the distance effect and the size effect. This is what affects the ability to identify the larger of two numerical quantities or tell whether two numerical quantities are the same (Dehaene, 2001). It is believed that the distance and size effects are also represented in symbolic notion as Arabic numerals and number words. From this, it is concluded that the adult brain contains an analogue representation of numerical quantities, similar to those in animals and infants. Additionally, when adults encounter number words or Arabic numerals, they convert them from symbolic format to the analogue quantity representation (Dehaene, 2001).

If NS is an evolutionary trait, there should be some evidence in the brain. There are two arguments that support this. The first is that the internal representation of quantities can be impaired by a brain lesion. The second is that brain-imaging techniques must show specific areas with activity during number processing tasks (Dehaene, 2001). Acalculia (*acquired* dyscalculia) shows the same systematic fall in the inferior parietal region across generations and cultures. This is the same area that has been reported for those with a deficit in number processing, also known as *developmental* dyscalculia (Dehaene, 2001). It is believed that both hemispheres of the brain, presumably, are involved in the analog quantity representation. This indicates that a deficit in NS, may cause the preverbal subtype of DD.

NS is often referred to as the ability to subitize, count, discriminate quantities, discern number patterns, exclude unreasonable results of arithmetic operations, move between different numerical formats and the approximate number system (ANS). There are different views on the ANS, and we'll look further into it below.

2.4.1 Approximate number sense

ANS is present at birth, have been shown in both preverbal infants and animals, and ANS is independent of language and culturally bound number symbols (Agrillo, 2015; Libertus et al., 2011). The ANS is active throughout life, from infancy into old age, and functions as a support for the primitive sense of numbers (Libertus et al., 2011). ANS has been identified to reside in the intraparietal sulcus of the parietal lobe in the brain (Neider & Dehaene, 2009).

ANS is an imprecise ability that estimates and mentally combines approximate quantities of objects. In other words, it is a fast, inherent, intuitive ability to approximate, i.e., assess quantities without counting (Bonny & Lourenco, 2013; Libertus et al., 2011; Mazzocco et al., 2011). Infants, children and adults can compare, add and subtract sets of dot arrays or sound sequences. ANS follows Weber's law, which dictates that the extent to which two quantities can be discriminated is determined by their ratio. This means that ANS is the minimum change in quantity that an individual can discriminate; it is the smallest 'just noticeable difference'. An individual's acuity of ANS is defined by the ratio at which the individual can reliably discriminate between two non-symbolic quantities (Bonny & Lourenco, 2013; Libertus et al., 2011; Mazzocco et al., 2011; Sasanguie, Göbel, Moll, Smets

& Reynovet, 2013). The acuity of ANS has been shown to increase in precision over the life span. When discriminating non-symbolic visual arrays, newborn is able to discriminate at a ratio of 1:3, whereas 6-month-olds at ratio of 1:2 (Bonny & Lourenco, 2013). A study by Halberda and Feigenson (2008) found that ANS does not attain full acuity until later during development. Their results show that 3-yer-olds can discriminate at a ratio of 3:4, 4-year-olds at 4:5, 5- and 6-years-olds at 5:6 and adults 10:11. This indicate that maturation is a factor in increasing ANS's acuity. Other studies have also found that experience and education are factors that enhance ANS acuity (Piazza, Pica, Izard, Spelke, & Dehaene, 2013; Hyde, Khanum, & Spelke, 2014). According to findings, an individual's ANS acuity is related to their performance in math tests (Mazzocco et al., 2011).

As mention, ANS acuity follows Weber's law. ANS acuity is often measured with a dot-comparison task. The participant is presented with to two quantities of dots (Dehaene, 2007) and is told to mark the most numerous amounts by estimating as quickly as they can. If the task is done on paper, the participant often gets 30 seconds to mark as many items as possible. When conducted on a computer, the participant is placed at a specific distance from the screen and gets to push buttons to mark the largest quantity. Outcomes of ANS acuity tasks can be accuracy, response time (RT) or Weber fraction (*w).* Accuracy is usually the percentage of correct responses, and RT is the time used from image onset to key press (Libertus, et a., 2011). Weber fraction (*w)* is often measured using a psychophysical model or sigmoid model. The Weber fraction *w* represents the amount of noise in a participant's ANS acuity (Halberda & Feigenson, 2008). When measuring ANS acuity at several time points, it is common to observe that accuracy increases while RT decreases. Also, that participant's accuracy decreases when numerical ratio increases. This resulting in a lower *w* score, indicating higher accuracy (Libertus, Feigenson, & Halberda, 2013).

ANS acuity develops over time, and when children learn to count and learn formal mathematics they acquire a symbolic system for representing numbers. The new system gives a precise representation of quantity that can be compared and manipulated. The non-symbolic system and the symbolic system are then mapped onto one another (Mundy & Gilmore, 2009). In the same way that ANS acuity follows Weber's law, the speed of comparing symbolic representations is also affected by the distance and the ratio between the digits. This effect, of the symbolic representation is a result of being mapped onto the underlying non-symbolic representations. This means that the non-symbolic system (ANS), affects the individual's ability to compare symbolic numerals. Mundy and Gilmore (2009) found in their study that children’s ability to map between non-symbolic and symbolic representations is related to the children’s mathematical achievements. Furthermore, they found that children’s accuracy in symbolic and non-symbolic comparison abilities were related to their mathematical performance.

2.5 Approximate number sense and arithmetic abilities

As of today, multiple studies have been performed on the relationship between ANS and arithmetic abilities, and there are major disagreements in the scientific community. It is suggested that ANS forms the basis for the development of arithmetic skills, before learning the Arabic numerals (Sasanguie et al., 2013). Furthermore, ANS makes sense of numerals, when they are presented (Mundy & Gilmore, 2009). Some studies have found that there is a correlation between children’s ANS and their later math performances, but how this correlation can be explained remains uncertain (Libertus et al., 2011; DeWind & Brannon, 2013; Halberda & Feigenson, 2008). Could a deficit in ANS had resulted in impaired mathematical development, or have the mathematical development, resulted in a poor development of ANS? Other studies have concluded that ANS does not play a significant role in math abilities. Instead they claim that the main factor is early numeracy (Brannon & Park, 2016). We will now view some studies that have investigated the relation between ANS and mathematical ability, with an emphasis on arithmetic. Discussion of these studies will be grouped based on their research design.

2.5.1 Cross-sectional studies

Cross-sectional studies contain data from one time point and are used to study the relation between different factors (Field, 2009; Gall, Gall, & Borg, 2007). For example, Libertus, Feigenson and Halberda (2011) aimed to examine the relationship between ANS acuity and early math abilities in young children before they received formal mathematic instructions, while controlling for non-mathematical abilities. They tested 200 3- to 5-year-old ($M_{age} = 4.2$ *years*) children’s non-symbolic comparison skills and their math abilities and made an assessment of verbal ability. They used a dot comparison task performed on a computer, which measured the children’s ANS accuracy, Weber fraction and response time (RT). To assess the children’s early mathematical ability, they tested their knowledge of verbal counting with one-to-one correspondence, number comparison skills, numerical literacy, number facts retrieval, calculation skills and number concepts. Children's expressive vocabulary was measured by parental report, telling how many words they had heard the child say. In the analysis, it was tested to which extent accuracy, Weber fraction and RT of ANS correlated with math ability. They found that each of the measurements significantly correlated with math ability. When they controlled for age and vocabulary size, they still found that accuracy, *w* and RT were unique contributors to math abilities.

Bonny and Lourenco (2013) aimed to examine the relation between ANS acuity and math competence, while dividing the children into groups based on age (3-, 4-, and 5-years-olds) and math abilities. Previously, studies have found a relation between ANS acuity and math abilities in young children, but they have not examined development within age groups. Bonny and Lourenco (2013) now wanted to examine development and the relation for children from ages 3 to 5, to see if development is linear or not. To test ANS acuity they used a dot-comparison task, measuring accuracy. To test math abilities, children's verbal counting, cardinality and elementary numerical arithmetic abilities were assessed. In addition, they tested the children's receptive vocabulary. First, they found that ANS acuity develops, and that older children performed better than younger children. In addition, as according to Weber's law, accuracy depended on ratio. They also found that ANS predicted math abilities while controlling for receptive vocabulary and age, across the age range. Moreover, they found a positive relation between ANS acuity and math abilities, indicating that children with more accurate ANS performs better in maths. When they assessed the relation for each age group separately, the finding was not consistent. The result showed that there was a marginally significant correlation between ANS acuity and math abilities for 3-yer-olds, while there was a significant relation for 4-year-olds, and no significant relation for 5-year-olds. This suggest that the relation between ANS and symbolic mathematics is nonlinear.

Fuhs and McNeil (2013) examined, on the grounds of recent findings, if the relation between ANS acuity and math performance would hold when testing ANS and math abilities of pre-schoolers from low-income homes. They hypothesised that the relation between ANS acuity and mathematic abilities is weaker in a sample of children from low-income homes, because children in their study have less exposure to math in everyday life. In addition, they hypothesised that inhibitory control would be significantly correlated with ANS acuity and mathematical abilities. The sample consisted of 103 pre-school children from age 3 to 6 ($M_{age} = 4.5$ *years*) from low-income homes. ANS acuity was measured with an objectivecomparison task on paper. Mathematical abilities were tested with a standardised test (TEMA-3) assessing relative magnitude, counting skills, number and numerical calculation skills, number facts and base 10 concepts. Inhibitory control (i.e., the ability to suppress a premature response) was tested by Head/Feet; the experimenter said ‘feet’, and the child had to touch their head instead of feet, and similar tasks. Receptive vocabulary was also tested. They found that the accuracy of ANS acuity was marginally related to mathematical abilities, and that the relation disappeared when controlling for receptive vocabulary. They also found that inhibitory control was significantly correlated with ANS acuity and mathematical abilities. These results indicate that inhibitory control might have a role in the relation between ANS acuity and early mathematical abilities.

The findings from the cross-sectional studies are conflicting, even though they all have assessed ANS’s relation to math abilities within the same age range (3- to 6-year-olds). At this point it is not clear if it should be expected to find a relation between ANS and arithmetic abilities for the 5-years-olds in the sample of the current study. Further, it is not clear that such a relation has a predictive role at this stage.

2.5.2 Longitudinal studies

Longitudinal studies collect data from multiple timepoints across a specific period. It is a valuable design when a study aims to describe change and development in a sample. Longitudinal studies are the farthest we can go without testing causes directly but can therefore also not say anything about causal relation (Gall et al., 2007).

Mazzocco, Feigenson and Halberda (2011) assessed if early ANS acuity predicts mathematical abilities during or after kindergarten, when controlling for general full-scale IQ. They measured accuracy of ANS acuity by letting children discriminate between quantities of identical and familiar objects. The sample contained 17 3-to 4-year-old children at the first time point, who were retested two years later. They found that ANS in 3- and 4- year-olds was a significant predicter of mathematical abilities two years later, and that the association could not be explained by full-scale IQ.

Libertus, Feigenson and Halberda (2013) examined if early ANS acuity at age 4 has a predictive role for math abilities 6 months later. ANS was measured with both accuracy and reaction time (RT), while controlling for individual differences in mathematical ability at the first measuring point. They also controlled for expressive vocabulary, attention and working memory. Attention and working memory were measured to see if earlier findings of ANS acuity and its relation to math abilities, is a result of ANS-specific influence or could be attributed to domain-general cognitive skills. The sample consisted of 204 preschool children (M_{age}*=4.2 years*), measured twice with a 6-month delay. They found that accuracy and RT for ANS acuity tasks were significant unique predictors of later math ability, even when controlling for initial math ability, age and expressive vocabulary. Second, they found that accuracy and RT of ANS acuity were unique predictors of math ability, when controlling for attention, memory span and expressive vocabulary.

Sasanguie, Göbel, Moll, Smets and Reynvoet (2013) examined if it was accurate ANS, symbolic number comparison or number-space mapping that is the best predictor of future mathematical abilities. They tested 92 children from the 1st to the 3rd grade and again one year later. At the first time point, they tested symbolic and non-symbolic number line estimation as well as symbolic and non-symbolic comparison skills, in addition to a general curriculum maths test. At the second timepoint, they administered a general curriculum maths test and a timed arithmetic test. They found that ANS at age 6 to 8 could not explain mathematical achievements one year later. Instead, they found that symbolic (number) comparison is the main predictor for children's mathematical achievements one year later. They conclude that mathematical difficulties are a result of a delayed response when assessing symbolic magnitude.

Desoete, Ceulemans, De Weerdt and Pieters (2012) examined the predictive value of individual differences in symbolic and non-symbolic comparison skills for arithmetical achievement. In addition, they looked at the development of comparison skills from kindergarten, and two years later. They also examined children's symbolic number and non-symbolic comparison abilities of those scoring under the 25th (LA) and the 10th (MD/DD) percentiles for arithmetical achievements to see if there is a predictive relation between these skills, controlling for intelligence. The sample consisted of 395 children. They were tested in Kindergarten (age 5 to 6) and again in Grade 2 (age 7 to 8). They organised the testing in line with the triple code model; ANS acuity was tested with accuracy (analogue magnitude code), symbolic comparison using number words (auditory verbal code) and Arabic numerals (visual Arabic code). They found no significant difference in intelligence between those with MD, LA and typical achieving children. They found that non-symbolic skills in kindergarten predict overall arithmetical achievement one year later, but only arithmetic fact retrieval two years later. Arabic numeral comparison skill predicted calculation abilities two years later. In addition, they found that children with MD had a significant deficit in their accuracy of ANS acuity and Arabic numerals in kindergarten, and children with LA had a mild problem (that they performed poorer than typical achievers in Kindergarten, but better than children with MD) with the ANS acuity comparison task concerning accuracy. These results indicate that a combined deficit in ANS and symbolic numeracy represents a risk for developing MD.

The main aim of Göbel, Watson, Lervåg and Humle (2014) investigated whether tasks used to measure ANS are reliable. In addition, they aimed to identify longitudinal predictors of arithmetic. They assessed children's non-verbal ability, vocabulary knowledge, letter comparison ability, magnitude-comparison (digit and dots) and arithmetic skills. The sample consisted of 173 6-year-old children, who were measured twice with an 11-month gap. First, they found that non-symbolic comparison tasks are reliable, and do measure ANS ability. Moreover, they found that ANS, letter comparison and number-identification were strongly correlated with arithmetic skills, but only number-identification was a unique predictor of arithmetic skills in children at this age.

Toll, van Viersen, Kroesbergen and van Luit (2015) aimed to, 1) examine the developmental relation between non-symbolic and symbolic comparison skills over time, and 2) assess how growth in non-symbolic and symbolic comparison skills can predict mapping and other basic mathematical abilities (i.e., number, number relation and simple addition and subtraction tasks verbalized by the teacher and arithmetic fluency tasks). To answer this, 671 kindergarten children with a mean age of 4,6 years participated in a longitudinal study covering two and a half years, from the first year of Kindergarten to the end of 1st grade. During the Kindergarten years non-symbolic and symbolic comparison skills were measured four times. Their intelligence was measured halfway through 1st grade and their basic arithmetic and mathematics and mapping was measured at the at the end of 1st grade. A dot-comparison task was used to measure non-symbolic comparison skills, and a digit-comparison task used to measure symbolic comparison skills. To measure mapping skills, a (symbolic) number line task was used. Mathematical reasoning was measured with combination of word problems, number relations and simple addition and subtraction calculations. Arithmetic fluency was tested by asking the children to complete as many addition and subtraction tasks as possible within one minute. Their results showed that non-symbolic and symbolic comparison skills developed significantly during kindergarten, both influencing each other mutually. In addition, they found that the growth in non-symbolic comparison skills was significantly related to development in symbolic comparison skills. Finally, they found that symbolic comparison skills were the main predictor of children's mapping skills, math fluency and math reasoning ability at end of first grade, while non-symbolic comparison skills only predicted math fluency levels at the end of first grade.

Overall the longitudinal study discloses that there is a relation between ANS and mathematical and arithmetic abilities within the different age groups. So, at this point to expect a relation between ANS at age 5 and arithmetic abilities at age 6 in the sample of the current study are justified. Though, the question if ANS has a predictive role for arithmetic, especially for the age group in this study is unclear.

2.5.3 Experimental studies

Experimental designs is very suitable when a study aims to find causal relationships between two or more variables. In general, the independent variable is manipulated, either by an intervention or within an experimental setting to measure its effect on the dependent variable (Gall, et al., 2007).

Obersteiner, Reiss and Ufer (2013) studied whether training exact (subitizing) or approximate (ANS) number sense would affect achievement in arithmetic. They wanted to see if there are different effects from training exact or approximate number sense on basic number processing skills in 1st grade, and if there are different effects on arithmetic achievement. With a 2x2 design and pre- and posttest measurements, they divided 204 first graders into four groups. A group trained in approximate number sense, a group trained in exact number sense, a group trained in both skills and a control group. Training of the exact and approximate number sense was done by means of computer games. Exact skills were trained using organized dot patterns to enhance exact mental quantity representation, and approximation by using random dot patterns and analogue representations to enhance approximate skills were trained using larger quantity random dot patterns and analogue representations, to enhance approximate mental quantity representations. The control group used language software. All children were pre- and posttested for basic number processing (exact and approximate) and arithmetic abilities (number sequences, number ordering, addition/subtraction, number line). They found that the approximate training only supported tasks where approximate mental representation were needed and the exact training only had positive effects on conceptual subitizing task. The combined training had no significant effect on arithmetic tasks. Finally, receiving training of exact number sense or training of approximate number sense can result in improving achievements in arithmetic.

Piazza, Pica, Izard, Spelke and Dehaene (2013) aimed to clarify whether it is maturation or education that affects the development of ANS acuity. To test this, their sample consisted of an indigenous population from the Amazon region, the members of which had varying access to education. As a result, many of the participants have a restricted lexicon for number words and no symbolic system for exact numbers and arithmetic. Hence, if maturation causes development of ANS acuity, then educational level should not affect the ANS acuity. In contrast, if education affects ANS acuity, then the more educated participants should have a more refined ANS. They did two experiments to test this. The main experiment contained 38 participants between the age of 4 and 63 that had zero or more than four years of schooling (children are generally first presented with numbers and basic arithmetic in 3rd school year in this population). ANS in this experiment was tested by dot comparison tasks on computer. The second experiment contained 33 participants between 4 and 67 years, whit the same level of education as those in the first experiment. Here ANS was measured with a comparison task conducted on computer. They found that those who have had some education had a significantly more refined ANS acuity when controlling for age, and most significantly, a reduction of the Weber fraction was observed in the participants who had received counting and arithmetic schooling. Therefore, they concluded that education plays a significant role in the development to ANS acuity.

Hyde, Khanum and Spelke (2014) conducted two experiments based on the claim that activation of ANS may enhance children's performance in symbolic arithmetic. In the first experiment, 96 first graders participated. They were divided into four training groups. Each group received training in a particular non-symbolic magnitude skill, expected to play a role in symbolic mathematics- non-symbolic numerical addition, line length addition, non-symbolic numerical comparison and brightness comparison (comparing what image is brightest). All training was conducted on computers, and the participants did arithmetic tasks on paper, before and immediately after the training sessions, consisting of 60 problems. First, they found, that children from the numerical addition and numerical comparison training groups were faster at completing arithmetic problems than those in the two other groups. After checking for improvement in accuracy on the arithmetic task, it turned out to be the non-symbolic numerical addition group that showed the greatest improvement. This led to the conclusion that numerical addition is most suitable for training ANS for improving symbolic mathematical skills.

To test if the number addition practice from Experiment 1 was specific to mathematics they hypothesized that the same training, should not affect sentence completion problems. To test this, in the second experiment, they again administered the non-symbolic approximate addition task training, and a similar sentence completion test, where the children were presented with a sentence and should complete this with a supplied word. The sample consisted of 48 first graders divided into two groups; non-symbolic numerical addition training and brightness comparison training. The children were given either the sentence completion test or symbolic arithmetic problems during the training. After half of the training they received the other one. The result showed, consistent with the first experiment, that the numerical addition group solved arithmetic problems quicker than the brightness group. There was no difference between the groups in solving sentence completion test problems. Furthermore, the non-symbolic approximate addition training performed more accurately on the arithmetic problems than those in the brightness group. The improvement was also limited to problems in the domain of mathematics. Based on the two experiments, there is some indication that there is overlap between the neural structures used for symbolic arithmetic and ANS.

Wang, Odic, Halberda and Feigenson (2016) examined whether a temporary modulation of ANS precision changes symbolic math performance, testing for a causal link between ANS and symbolic math. Firstly, they wanted to test if children do more precise ANS discriminations if they meet easy tasks first then gradual progress to harder once. Therefore, in Experiment 1 a sample of 40 5-year-olds were divided into two groups, one received increasingly difficult tasks, the other decreasingly difficult. Half of the children in each group were tested on a symbolic math task, and the other half on a non-numerical vocabulary task. ANS acuity was tested with a dot-comparison task conducted on computer. Symbolic math abilities were tested with symbolic math, such as verbal counting, arithmetic tasks, knowledge of Arabic numerals, and understanding of place value. They also assessed the children's vocabulary. Firstly, they found that children's ANS discrimination can be altered. Secondly, they found that the changes in the children's ANS (enhanced or impaired) transferred to children's symbolic math achievement, but not to their vocabulary knowledge. This provides proof of a causal link between ANS and symbolic math, but it is not clear if children's math performance was improved, disrupted or both. Therefore, in a supplementary experiment they conducted the same ANS discrimination task, but this time the degree of difficulty was randomized. The supplementary sample consisted of 10 5-year-olds. They found that ANS acuity in the random group improved at an intermediate level, between easy first and hard first groups. Suggesting that ANS can be both enhanced and disrupted. This result also appeared when concerning the result of the mathematical test, again suggesting a causal link between ANS acuity and symbolic math performance.

2.6 Empirical findings

In 1992 Dehaene developed a simplified framework for number processing. Based on this model, there is a primitive inherent skill that lays the foundation for analogue magnitude perception. This ability constitutes the first building block for the ability to acquire other mathematical knowledge such as number words, Arabic numerals and arithmetic. However, whether this view of mathematical development is correct is still debated. This theoretical basis for TCM fits with other theories such as biological primary and secondary skills (Geary, 2000) as it divides mathematical abilities into abilities that we are born with and abilities that we learn while growing up. Furthermore, the TCM model fits with the view on general-domain and specific-domain precursors (Passolunghi & Lanfranchi, 2012). That divides mathematical abilities into specific abilities that are required within the mathematical domain, and other abilities that also necessary, but that are important on a more general basis to learn and function. Based on the TCM a deficit in one of the three modules can lead to DD. This can explain why DD and LA are expressed in different ways in different individuals, and why it can occur at different stages during development. As the development of these abilities build upon each other, while the development of a new ability strengthens the previous ability (Von Aster, 2000; Von Aster & Shalev, 2007).

The empirical findings of previous studies show an academic field with uncertainties, different results and theoretical explanations for why, or why not, NS and ANS affect arithmetic abilities. Some studies found that ANS has a strong relation with arithmetic (Bonny & Lourenco, 2013; Libertus et al., 2011), and mathematical abilities (Desoete et al., 2012; Libertus et a., 2013; Mazzocco et al., 2011; Toll, et al., 2015). A recurring factor of these studies is that the first test point is performed with children at an early age, between 3- and 5-years-old. On the other hand, there are studies that find that ANS has only a small relation with arithmetic, that disappears when it is controlled for vocabulary and early numeracy (Fuhs & McNeil, 2013; Göbel et al., 2014; Sasanguie et al., 2014). In two of these, this is the case when early numeracy and symbolic abilities are the main predictors of arithmetic. But in contrast to the other studies, the children in this case are of an older age group, from age 6 and up. Experimental studies have shown that ANS can be improved, but also disrupted, training has resulted in either increased or decreased arithmetic achievement. Furthermore, these studies found that ANS overlaps with arithmetic, and that the training has a transfer effect from ANS to symbolic numeracy. They have also shown that education is an important factor for how refined one's ANS acuity (Hyde et al., 2014; Obersteiner et al., 2013; Park & Brannon, 2013; Piazza et al., 2013; Wang et al., 2016).

To further extend the literature on this topic, the current study will address the relation between ANS and arithmetic in 5-year-olds Norwegian kindergartners. It will assess 1) to what extent NS at age five can predict arithmetic ability halfway through 1st grade, and 2) whether this relation persist after controlling for intelligence, vocabulary and early numeracy. Based on the theory and empirical findings described above, several hypotheses have been formulated. First, ANS acuity at age 5 can predict ANS acuity one year later, since the same concept are measured at both timepoints. Second, it is expected that ANS in-5-years-olds will predict arithmetic skills on year later. This is because the children are tested before they start school and have received limited formal mathematical instructions in kindergarten. Consequently, it is also, hypnotized that the relation between ANS and arithmetic will remain present after controlling for non-verbal intelligence, vocabulary and early numeracy.

3 Method

This chapter presents the methodological approach of the study. The study design and participant selection are described, as well as the procedure for collecting data and the measuring instruments. The validity and the reliability are also addressed, as well as ethical considerations.

3.1 Design

The aim of the study is to investigate the extent to which there is a relationship between NS in 5-year-olds and their arithmetic skills after the first semester in 1st grade, while early numeracy is accounted for. The study is performed using a quantitative approach with a non-experimental observational longitudinal design. The study is descriptive, as it tries to describe reality without affecting or manipulating variables through interventions (Kleven, 2002b). Data from two measurement points, at age 5 in Kindergarten and age 6 in 1st grade is used for the analyses.

3.2 Participants

This study is part of the longitudinal research project *Development of Numeracy and Literacy in Children (NumLit)* at the Department of Special Needs Education of the University of Oslo, which follows children from the age of 5 (Kindergarten) until they are 18-year-old (Grade 13). For the current study, the sample was limited to Norwegian monolingual children without major learning difficulties. The sample consisted of 170 pupils, 49.1 % boys with a mean age of 5.49 years at the first timepoint. The pupils came from Lørenskog municipality, Skedsmo municipality and Oppegård municipality. The sample is considered to be representative of the Norwegian population in relation to parental education and socio-economic background. NumLit is registered and approved by the Norwegian Centre for Research Data (NSD) Therefore, the study is carried out in accordance with the ethical guidelines *The National Research Ethics Committee for Social Sciences and Humanities* (NESH, 2016). Since data on children was collected, informed consent was asked and provided by all parents as they agreed on behalf of the child. Parents of children in the target group were asked to participate by the kindergarten staff. Consent forms were sent to the parents via the kindergartens. Further ethical considerations are elaborated in Part 3.6.

3.3 Measurement instruments

3.3.1 Non-verbal intelligence

Non-verbal intelligence was measured using *The Raven Colored Progressive Matrices* (Raven, 1998). It gives a measure of the child's general cognitive abilities, and its administration does not require any language skills. The test was conducted on a computer. The child received a minimum amount of instruction and had two practice trails. The child would see a square that is missing a piece and underneath the square are six suggestions of patterns that could fit in-to the missing space in the square. The child should choose the piece that he or she thinks fits. There is no time limit and it is programed to stop by itself after a specific number of errors. The test has 34 items, and one point is awarded for every correct answer. Reliability of the task within the current sample is good, with Cronbach's α over .73.

3.3.2 Receptive vocabulary

Vocabulary was measured using the second version of the *British Picture Vocabulary Scale* (BPVS II), which is similar to the American PPVT (Dunn, Dunn, Whetton, & Burley, 1997). BPVS tests receptive vocabulary and has been translated into Norwegian and normed and standardized for the Norwegian population (Lyster, Horn, & Rygvold, 2010). The test was conducted on a computer and was especially made for the NumLit project by Athanassios Protopapas. The electronic version is fully congruent with the analogue version. The child hears a word and is presented with four pictures, from which the child chooses the picture that corresponds with the word. There is no time limit and the test are terminated after a specific number of wrong answers. The test consists of 144 items and one point is awarded for every correct answer. Reliability of the task within the current sample is good, with a Cronbach's α over .89.

3.3.3 Number sense

Number sense was measured with *Dot Comparison*, a subtest from TOBANS (Brigestocke, Moll, & Wiebe, 2016). This subtest gives an indication of approximate quantity estimation. The pupil assesses two boxes with different amounts of dots, where the pupil should mark the box representing the largest quantity. The pupil first completes three practice items and then completes the test, with a 30 second limitation. Task difficulty increases gradually, progressing from easy discrimination to harder discrimination. The test is performed on paper. The maximum score is 52, where one point is given for each correct answer. Raw accuracy scores are used. Reliability of the task within the current sample is good, with a Cronbach's α over .82.

3.3.4 Early numeracy

Early numeracy was measured with two experimental tasks, specifically made for the NumLit project. The first task is *Number Reading*. It measures the child's knowledge of Arabic numerals. The pupil is asked to read numbers aloud. It starts with low numbers and rises as the child progresses. The test consists of 4 items. Reliability of the task within the current sample is good, with a Cronbach's α over .84.

The second task is *Count On,* which measures the child's knowledge of whole-number-word sequences. It is a test where the pupil hears the test leader say three numbers, and then asks for the next number in the sequence. The test leader says for example "three, four, five, and then comes?" The difficulty level increases throughout the test. The test stops after four wrong answers in a row. The test consists of 8 items, which is also the maximum score. Reliability of this task within the current sample is good, with a Cronbach's α over .85

For the analysis, scores on both tasks are *z*-transformed and combined into one variable to measure early numeracy.

3.3.5 Arithmetic Addition fluency

Arithmetic skills were measured using two subtests from TOBANS (Brigstocke et al., 2016), an addition fluency test and a subtraction fluency test. These subtests measure the pupil's arithmetic fluency and accuracy skills for addition and subtraction sums. The pupil is instructed to make as many sums as possible within a time limit of one minute per subtest. The child is not informed about the type of sums before the subtest starts. The child gets three practice trials. The test is conducted on paper. The test consists of 60 items per subtest, and there is given one point for each correct answer. Reliability of the tasks within the current sample is good, with a Cronbach's α over .87 for addition and over .88 for subtraction, respectively. Raw accuracy scores on both subtests are combined into a total arithmetic score for the analyses.

3.4 Procedure

For the data included in the present study, children were tested on different abilities on three occasions, between December 2017 and March 2018 for the first time point (Kindergarten), and three times between December 2018 and April 2019 (1st grade) for the second time point. The test battery used contains established, standardized (sub)tests, as well as test materials specifically developed for the project. The tests were administered by approximately 20 research assistants at both time points, but not necessarily the same ones. The research assistants had been trained in administering the tests with group meetings by the researchers of the NumLit project employed by the University of Oslo. Each test session lasted between 45 and 70 minutes, and as far as possible, the same research assistant tested the same child during consecutive measurements. The tests were carried out in a quiet room, with an audio recorder. The children received stickers for each completed assignment and a diploma at the end of each test round. The data is only available for the study. Parents and schools were not given information about the results.

3.5 Validity and reliability assessment

To assess the credibility of the findings in this study, the validity system made by Cook and Campbell (1979) is used. The validity system is intended for research designs that study causal relationships but is also used for non-experimental studies. In this part, four types of validity are discussed; construct validity, statistical validity, internal validity and external validity. Reliability is also addressed. The validity and reliability and their importance for this study are further discussed in Part 5.2, to assess the degree to which the results of this study can be generalized.

3.5.1 Construct validity

Construct validity is the extent to which the concepts to be investigated is actually measured with the given measuring instrument. Concepts that are investigated within the educational and psychosocial research field are often not directly measurables. Therefore, theoretical concepts must be operationalized so that they become measurable (Field, 2009; Kleven, 2002a). In this study, construct validity is about how well the theoretical variables; number sense, early numeracy, non-verbal intelligence, vocabulary and arithmetic skills are actually measured by the described instruments (Kleven, 2002a). The test reliabilities are important for the validity. However, a test can have high reliability, but low validity when it does not the measure intended concept. Therefore, the term used for the operatized concepts are important for the concept validity (Lund, 2002) and construct validity has therefore an important role in my study.

3.5.2 Statistical validity

Statistical validity is about the extent to which there is covariation between the independent and the dependent variable, and how strong the covariation is (Shadish, et al., 2002). To tell if the covariation is of theoretical importance, there must be strong statistical significance (Lund, 2002; Shadish, et al., 2002). With good statistical validity, is it less likely to make type I and type II errors. Type I errors occur when one rejects the null hypothesis, concluding that there is a relation between two variables, while there is actually no relation. A type-II error is when one concludes that there is no relation and accepts the null hypothesis, while there actually is a relation. Reliability is also important in statistical validity. Random measurement errors during the data collection, may result in weak reliability, which results in incorrect conclusions about the significance of results (Shadish, et al., 2002).

3.5.3 Internal validity

Internal validity is whether the observed relationship between variables refers to a causal relationship - if the causal relationship is valid based on the data analysis (Shadish et al., 2002). Choice of design and method is important; non-experimental designs lead to a weak internal validity, as causal conclusions cannot be drawn. That is the directional statistical problem; ‘which variable leads to the other cannot be answered (Lund, 20002). In this case, previous research becomes a central part of being able to interpret the results, and in being able to arrive at causal hypotheses based on previous findings. Another threat can come from the measurement instrument or the procedure. For example, ceiling and/or floor effects can be a threat to the internal validity, which can affect the degree to which the results can be generalized. Thus, internal validity is of interest in this study.

3.5.4 External validity

The possibility to generalize research is highly relevant in educational science. The extent to which the results can generalized says something about the external validity (Shadish et al., 2002). Generalization means transferring the result from a smaller sample to a larger population (Lund, 2002). Threats to the external validity may be a non-representative sample, i.e., the sample does not represent the relevant characteristics of the population. Furthermore, low statistical validity will pose a threat, as the ability to draw valid conclusions is impaired. This will result in a low transfer to the population (Lund, 2002). Therefore, the external validity will be discussed in comparison, to the sample in this study and the degree to which the data meet the assumptions that are required to conduct the analysis.

3.5.5 Reliability

In order to answer the research question, the variables were selected based on earlier studies. The selected tests are similar or identical to tests used in other studies to measure these abilities. In this way, the results can be compared with other studies that have used the same measurement instruments. To evaluate reliability, Cronbach's alpha was used to measure intercorrelations among test items within the same instrument. Cronbach's alpha should be .70 or higher for the instrument to be evaluated as reliable (De Vaus, 2014).

3.6 Ethics

The National Research Ethics Committee for Social Sciences and Humanities (NESH; 2016) provides guidelines for safeguarding good research ethics in *Research Ethics Guidelines for Social Sciences, Humanities Law and Theology*. Research ethics refers to values, norms and institutional arrangements for the exercise of good scientific practice, with the need to safeguard the human equality and inviolability, as the foundation for privacy and the dignity of the individual(NESH, 2016). In this, the Universal Declaration of Human Rights (De Forente Nasjoner, 1989) and the Convention of Rights of the Child (De Forente Nasjoner, 1989) are the guiding principles for safeguarding people and children.

The research project that this master thesis is a part of are registered and approved by the Norwegian Center for Research Data (NSD). They assess research projects with regard to personal data, data collection, data archiving and ethical considerations. This study is conducted in accordance with the ethical guidelines by NESH. Data security and privacy information are secured by internal routines at the University of Oslo. Since children are the basis for the data collection in the project, informed consent letters were given to all parents. The parents gave consent on behalf of the child. This lay the grounds for the study being ethically justifiable (Befring, 2015). The parents of the participating children were informed of the NumLit project through the kindergartens and an information letter. In this letter, the parents were informed about the aim of the study and how data would be handled to protect their and their child's privacy. In addition, they were informed about their rights to withdraw and in what instances to submit complaints to. Further, they were informed about who to contact if they had any concerns or questions. In the letter, the parents were also informed about how often and how long every test situation would be and it emphasized that the test situation should be a positive experience for the child. Participation in this project is voluntary; children and parents can withdraw at any time.

In order to safeguard the privacy, all personal data are treated confidentially. All researchers and research assistants on the project have a duty of confidentiality and all data is anonymized (NESH, 2016) by using an ID-number for each participant. The ID-numbers were used on the test protocols to ensure that data about the same pupil could be connected and combined. The results from children were not shared with parents, school or other parties outside the project to ensure that children could not be recognized at a current or later time point. Storage of collected data and personal information is cared for through the internal routine developed of the NumLit project. The data is stored on secure external hard drives and must be deleted by June 2032. Sound records of the test situations shall be deleted after the data have been coded, no later than 6 months after assessment took place. This is in line with the ethical guidelines of NESH (punkt 11, 2016) that state that personal information and collected data should not be stored longer than necessary to meet the study's purpose.

In research, children are considered as a particularly vulnerable group, because there is an asymmetric relationship between the pupil and the adult. This is because pupils lack the basis for the same insight and understanding as the adult. They don't necessarily see the consequences of participation in a project. Minimizing risk and burden to an insignificant level is therefore important (NESH, 2016). Because of the asymmetric relationship it can be difficult for the child to communicate their wish to and wish not to participate. It will then be important that the adult in the test situation lays the foundation for a good and safe environment for the pupil.

During the test situations, some pupils have stated that they had enjoyed themselves, while others have experienced the testing as more demanding. To ensure that the test situation is as predictable as possible of the pupils, they were informed by their teachers why they were being tested and how long it would take. Each test session took about 40 to 70 minutes. Before the test session started, the test leader tried to create a secure environment for the pupil, by telling about what was going to happened next, the child's right to stop when he or she wanted, and the test leader first started testing after the child had given a go ahead. If the pupil would express reluctance, it was particularly important that the test leader was motivational and supportive by giving the pupil positive feedback and acknowledging their feelings. Continuation would happen only on the pupil's premises to safeguard the pupil's integrity. The pupil had the opportunity to take breaks when needed and cancel single test or the entire test session. It is also a responsibility of the test leader to take breaks or interrupt if the test leader finds it necessary. The pupils were encouraged to do the best they could, but it was not important that they answered correctly, this to avoid performance pressure. During every test, the test leader gave process-based positive feedback regardless of the pupil's performance. For each completed test, the pupil received a sticker as a reward and for further motivation. The pupils liked this a lot and it worked as a natural break during testing.

There are many factors that can influence the collection of data. Such as the child's state of mind, a negative relationship between the pupil and the research assistant, and motivation. In this study there were many research assistants. They went to many schools in different municipalities, met principals and teachers, assistants, other employees and the pupils. The frame that is set before the actual testing can affect the pupil. During the data collection, various situations have arisen. For the most part, collecting data was a positive experience both for me as a research assistant, but also for the pupils that have participated. This assumption is based on personal reports from the pupils and their teachers. As a research assistant, teacher and special needs educator, some of the experience have raised questions about how the adults around the pupils might affect the pupil before the test session, and how this could have affected the result of the pupil. This is, for example, from experiencing that teachers or others referred to a pupil in a derogative way (e.g. "You can try to test him, but he is the worst of the worst" and "You will probably not manage to test him") Their statements were probably not thought through or intended for the pupils ears, but they still had the opportunity to hear what was said. Hearing this as a pupil right before being tested could have affected the pupil and might have become a self-fulfilling prophecy. This issue is further discussed in relation to validity in Part 5.2.4.

4 Results

In this chapter, the data are described and analyzed. First, screening of the data set and descriptive analysis of the variables are presented. Descriptive statistics are used to describe the sample characteristics for all relevant variables.

Further, the result section will describe and then explain the results. Bivariate correlations are used to summarize the relation between relevant variables. The correlation between variables that measure the same skills are first considered, to assess if they can be combined in-to one variable. Subsequently, a correlation analysis is performed on the newly defined variables to assess the strength of their relationships. The independent variables that are significantly related to the dependent variable arithmetic are included in the regression analysis. Multiple regression analysis is used to examine the strength of the predictive relationship between the independent variables and the dependent variable. The variables are entered in a hierarchical order. This enables assessment of the relation between number sense and arithmetic while controlling for early numeracy and several covariates. All analyses are done with the statistical program SPSS (IBM SPSS Statistics for Windows, Version 25.0).

4.1 Data Screening

4.1.1 Data cleaning

A descriptive analysis has been performed to describe the characteristics of the individual variables and to check assumptions for the final analyses. The total sample consisted of 187 children. Sixteen children were excluded because their Grade-1 data was not yet available. This reduced the sample size in the analysis to 171 children (49.1 % boys, $M_{age} = 5.49$). Remaining missing data was checked with a missing value analysis, using the Little's MCAR test. Overall, 0.74% of the data was missing. Little's MCAR test was not significant, indicating that the data were missing completely at random (MCAR). To examine univariate outliers, *z*-scores were computed for all variables. A *z*-score lower than -3.29 or higher than 3.29 was considered an outlier. Four univariate outliers were detected and adjusted to the closest raw score that was not considered an outlier: one outlier on the number sense task has been adjusted from 20 (*z*-score = 3.90) to 15, and two on vocabulary were adjusted from 7 and 9 (*z*-scores = -4.00 and -4.15), to 36 and 35. The last outlier was not changed, because it concerned a score that was near several other scores (i.e., 20, next score was 19). Multivariate outliers on the independent variables were checked using the Mahalanobis Distance with $p < .001$. There were no multivariate outliers.

Table 1 provides an overview of the variable characteristics after adjusting the outliers. Cronbach's alpha is a measure of the internal consistency of the instruments that were used to create these variables (i.e., reliability coefficient). The degree of inner consistency says something about how much scores on the individual items within tasks correlate with each other. Cronbach's alpha should have a value of at least .70 (De Vaus, 2014). Table 1 shows that Cronbach's alpha is higher than .70 on all tests, indicating that the tests have a good internal consistency and reliably measured the construct.

Table 1

Descriptive Statistics of the Variables after Data Cleaning (raw scores)

Variables	*N*	*M*	*SD*	Minimum	Maximum	Cronbach's *α*
Number sense						
Dot Comparison (K)	171	6.99	3.19	3.00	15.00	.83
Dot Comparison (G1)	171	10.74	3.22	0.00	20.00	.83
Early numeracy						
Number reading	171	8.26	2.78	1.00	13.00	.84
Count on	171	5.81	2.42	0.00	8.00	.85
Intelligence						
Raven	163	17.28	4.18	7.00	27.00	.76
Vocabulary						
BPVS	170	62.87	12.32	35.00	91.00	.89
Arithmetic						
Addition	171	9.63	3.50	3.00	21.00	.87
Subtraction	171	5.84	3.95	0.00	20.00	.89

Note. K = Kindergartner; G1 = Grad 1; BPVS = British Picture Vocabulary Scale

4.1.2 Checking assumptions

After adjustment of the outliers, the distribution of the variables was assessed to check for normality. Kolmogorov-Smirnov (K-S) tests were performed to check the shape of the distributions (see Table 2). A significant result ($p<.05$) indicates that the shape of the distribution for a specific variable based on the sample is significantly different from a normal distribution (Field, 2009). If results indicated that a variable's distribution deviated significantly from the normal distribution, the variable was further assessed using Skewness and Kurtosis indices. In a perfectly normal distribution, skewness and kurtosis are zero, but values within -1 to 1 are also considered good (de Vaus, 2014). A positive bias indicates that there are too many low values in the distribution, while a negative value indicates that there are too many high values.

Table 2

Kolmogorov-Smirnov test (K-S) and Skewness and Kurtosis Indices

		K-S			
Variables	*D*	*df*	*p*	Skewness	Kurtosis
Number sense					
Dot Comparison (K)	.09	163	.001	.117	.060
Dot Comparison (G1)	.90	163	.000	.300	.283
Early Numeracy					
Number reading	.10	163	.009	-.243	-422
Count on	.24	163	.000	-.935	.403
Intelligence					
Raven	.08	163	.003	.075	-.292
Vocabulary					
BPVS	.08	163	.000	.176	-.567
Arithmetic					
Addition	.14	163	.000	.776	.777
Subtraction	.100	163	.022	.636	.548

Note. K = Kindergartner; G1 = Grad 1; BPVS = British Picture Vocabulary Scale

Number sense is measured by *Dot Comparison* from TOBANS at two measurement occasions, Kindergarten (K) and 1st grade (G1). Both have alpha values over .82 indicating good internal consistency of the instrument at both time points. However, the K-S test reports a significant deviation from the normal distribution ($p < .001$ and $p < .003$). Nevertheless, dot comparison (K) is approximately normally distributed based on skewness and kurtosis indices (.117 and .060; see Figure 2, Appendix A). The distribution to dot comparison (G1) is also almost normal, with skewness and kurtosis at .300 and .283 (Figure 3, see Appendix A).

Early numeracy is measured with the tests ‘number reading’ and ‘count on’ specifically developed for the NumLit project. Both tasks have good alpha values over .84, which indicates good internal consistency. The K-S test reports a significant deviation from the normal distribution (p <. 001 and p <.001). However, the distribution of number reading is approximately normally distributed based on skewness and kurtosis at -.243 and -.422 (see also Figure 4, see Appendix A) indicating somewhat more higher values. The distribution of count on, is more heavily right skewed, suggesting a celling effect (see Figure 5, see Appendix A). However, as skewness is at -.935 and kurtosis is .403 and both values are still within -1 and 1, the distribution is acceptable, and the variable can be used further in the analysis.

Intelligence was measured with Raven (1998). Raven had an alpha value of .72, which indicates good internal consistency. The K-S test reports a significant deviation from the normal distribution (p <.009). However, the distribution of Raven is almost normally distributed, with a skewness and kurtosis of .075 and -.292 (see Figure 6, see Appendix A).

Vocabulary was measured with BPVS (II). BPVS had a high alpha value of .89, indicating a good internal consistency. The K-S test reports a significant deviation from the normal distribution (p <.022). The distribution of BPVS is approximately normal based on a skewness and kurtosis of .176 and -.567 (see Figure 7, see Appendix A).

Arithmetic was measured with the addition and subtraction tasks from TOBANS. Both tests had high alpha values over .87 that indicate good internal consistency. The K-S test reports a significant deviation from the normal distribution of both variables (p <.001 and p <.05). The distribution of addition is approximately normal based on skewness and kurtosis .776 and .777. (see Figure 8, see Appendix A) The distribution of subtraction has a slight left skew, indicating many low values and suggesting a floor effect (see Figure 9, see Appendix A). This can mean that the test was too hard for many of the pupils. However, skewness and kurtosis are of .636 and .548, indicating that the variable is approximately normally distributed.

4.2 Summary of assessment of the normal distribution

The descriptive analysis showed that some variables are not totally normally distributed, to some extent. This could have consequences for the interpretation of the bivariate correlation analysis. After closer evaluation of the skewness and kurtosis for each variable, the distributions can all be considered approximately normal. However, the test for counting on showed a celling effect and the test for subtraction fluency showed a floor effect. Since both of these variables are being combined into new variables with normally distributed variables, the skewness may disappear in the new variables. If the image of the combined variables (see below) changes the skewness, and the resulting variables can be considered approximately normally distributed, only Pearson's *r* will be reported for the correlation coefficient, as the assumption of normality is then met.

4.3 Combining variables

To test the relation between number sense, early numeracy and arithmetic, variables that measure largely the same abilities within these concepts were combined. Number sense was measured with a dot comparison task representing ANS acuity and therefore consisting of only one variable. Early numeracy was tested with the tasks *count on* and *number reading*. The tasks showed a strong and significant correlation (r = .57, p <.001) and were therefore standardized with *z*-scores and combined into an average score for early numeracy. For arithmetic, the subtests addition and subtraction were taken into account. The tasks correlated strongly (r = .70, p < .001) and were measured on the same scale. Therefore, they were combined into a sum score indicating arithmetic abilities.

After the combining of variables into two new variables, they were examined for new outliers, using the same methods as above. This resulted in one new univariate outlier for arithmetic, which was changed from 40 to 36 to make the distribution of arithmetic connected. In addition, one multivariate outlier was detected. Data for this child was excluded from further analyses, resulting in a final sample size of 170 children. Normality was also reassessed for the new variables (see Table 3).

The K-S test for the variable early numeracy reports a significant deviation from the normal distribution (p < .001). The distribution of early numeracy has a slight right skew, indicating somewhat more high values (see Figure 10, see Appendix B). This is likely an effect of the distribution to *count on*, that had a ceiling effect. However, skewness and kurtosis are at .450, and -.686, which are smaller values than for count on separately, indicating that the distribution can be considered approximately normal.

The K-S test for the variable arithmetic reports a significant deviation from the normal distribution ($p < .001$). When assessing the distribution of arithmetic, it seems a bit left skewed, indicating more low values (see Figure 11, see Appendix B). This probability comes from the subtraction task, that showed a floor effect. However, skewness and kurtosis of the combined variable are at .757 and .431, meeting the assumption for a normal distribution.

Table 3

Kolmogorov-Smirnov test (K-S) and Skewness and Kurtosis Indices

	K-S			Skewness	Kurtosis
Variables	*D*	*Df*	*p*		
Early Numeracy	.10	170	.001	-.450	-.686
Arithmetic	.98	170	.000	.757	.431

After creating the new variables, it was concluded that all of the variables meet the assumption of a normal distribution, with a skewness and kurtosis within -1 and 1. Therefore, the Pearson's *r* will be reported in the correlation analysis.

4.4 Correlation analysis

A bivariate correlation analysis was done to evaluate the relationship between arithmetic and number sense, early numeracy, intelligence, age and vocabulary. A bivariate correlation indicates how much the value of one variable changes as a function of changing the second variable. The correlation coefficient Pearson's *r* can have values between -1 and +1. Where values in the vicinity of 1 or - 1 indicate a nearly perfect correlation, a value of 0 or near 0 indicates no correlation (Field, 2009). The correlation coefficient says something about the strength of the relation and the direction of the association (i.e., positive or negative). A positive correlation means that if one variable increases, the other variable also increases. A negative correlation means that if one variable increases the other variable decreases (Befring, 2015). A correlation of .10 is considered weak, a correlation above .30 moderate, and a correlation above .50 is considered a strong correlation (Field, 2009). A two-tailed alpha-level of .05 is used to evaluate if the correlations are significant. Only variables that are significantly correlated with the outcome variables are taken into account in the regression analysis.

Table 4

Correlations between the variables

Variables	1.	2.	3.	4.	5.	6.	7.
1. Arithmetic	-	.					
2. Early Numeracy	.416**	-					
3. Number sense (K)	.231**	.304**	-				
4. Number sense (G1)	.914**	.400*	.210**	-			
5. Vocabulary	.107	.193*	.293**	.121	-		
6. Intelligence	.242**	.384**	.277**	.230**	.293**	-	
7. Age months	.190*	.184	.270**	.129	.172*	.102	-

Note. K = Kindergarten, G1 = Grad 1

* p <.05

** p < .01

The results of the correlation analysis show that there is a positive weak, but significant correlation between number sense in Kindergarten and in 1st grade. This is lower than expected because the same concept should be measured at both time points. This could indicate that the dot comparison task does not tap the same skill in Kindergarten and 1st grade. Number sense in 1st grade actually shows a very high positive correlation with arithmetic, suggesting that both variables may be targeting almost exactly the same skill. In contrast, Number sense (K) has a small positive correlation with arithmetic, and early numeracy a moderate positive correlation with arithmetic. Intelligence and age both have a weak positive correlation with arithmetic, while vocabulary is not associated with arithmetic. In line with these results number sense (K), early numeracy, intelligence and age will be taken into account as independent variables in the regression analysis, since they are significantly related to the dependent variable, arithmetic.

4.5 Hierarchical multiple regression analysis

4.5.1 Purposes

The bivariate correlation analysis indicated that there is a weak, but significant relation between number sense of Norwegian 5-year-olds and their arithmetic abilities at age 6. A hierarchical multiple regression analysis is therefore done to see to what extent number sense in kindergarten can predict arithmetic abilities in 1st grade, while taking age, non-verbal intelligence and early numeracy into account. A regression analysis gives additional information about how much variance in arithmetic can be explained by these factors.

4.5.2 Prerequisites for regression analysis

In social and psychological science, the ability to generalize from a sample to a wider population is of great value. To be able to draw these types of conclusion several underlying assumptions need to be met (Field, 2009).

Assumption for the outcome variable being an interval variable is met. In addition, all predictors have some variation in value, meeting the assumption for non-zero variance. The scatterplots for the distribution of residuals of predicted values show that the assumption of linearity is met (see Figure 14-17, Appendix C). The spread of the dots seems randomly and evenly spread. The same scatterplots show that the assumption of homoscedasticity is met. The assumption for multi-collinearity is met, none of the variables correlates more than .80 with another. Furthermore, all variance inflation factors (VIF) are below 10, ranging from 1.0 to 1.3. Durbin-Watson has a value of 1.773 and meets the assumption of independent residuals. The assumption for normally distributed residuals is met, and the variables raw score distributions are all approximately normal distributed (see Figure 12, Appendix C). In summary, all assumptions for conducting a regression analysis are met.

4.5.3 Hierarchical regression analysis

In general, the variables in a hierarchical regression analysis are selected based on previous research. Subsequently, the independent variables are entered in order of priority. This makes it possible to investigate whether new variables change the effect and increase explained variance by other variables in the analysis (Field, 2016). Here, a hierarchical multiple regression analysis was performed to examine the extent to which the independent variable number sense explains unique variance in arithmetic when age, non-verbal intelligence and early numeracy are taken into account. An additional goal was to determine which of the variables is the strongest predictor of arithmetic. The hierarchical multiple regression analysis is done in three steps, and the different models will be evaluated against each other. Table 5 and 6 give a summary of the hierarchical multiple regression analysis.

Table 5.

Hierarchical Multiple Regression Analysis

	Model 1			
Predictors	*B*	*SE*	95% CI	β
(Constant)	-13.22	9.19	[-31,38 – 4.94]	
Age	.34	.14	[0.07 – 0.62]	.187*
Raven	.36	.12	[0.12 – 0.61]	.23**
	Model 2			
Predictors	*B*	*SE*	95% CI	*β*
(Constant)	-10.13	9.32	[-28.533 – 8.275]	
Age	.28	.14	[0.01 – 0.56]	.15
Raven	.31	.13	[0.06 – 0.557]	.19*
Number sense (K)	.29	.17	[-0.05 – 0.64]	.14
	Model 3			
Predictors	*B*	*SE*	95% CI	*β*
(Constant)	-1.90	8.96	[-19.59 – 15.80]	
Age	.22	.14	[-0.05 – 0.48]	.12
Raven	.12	.13	[-.013 – 0.37]	.07
Number sense (K)	.15	.17	[-0.17 – 0.48]	.07
Early Numeracy	1.40	.31	[0.79 – 1.99]	.36***

Note.

* $p < .05$

** $p < .01$

*** $p < .001$

In Model 1, the influence of the covariates was assessed, and the independent variables age and non-verbal intelligence were entered. The first model shows that age and non-verbal intelligence are both significant predictors ($p < .05$) and combined explain 9.3% of the variance in arithmetic. Model 1 itself is significant (i.e., $F(2, 157) = 8.28$, $p < .090$).

In Model 2, number sense was added to the covariates as a predictor. The resulting model showed that number sense in Kindergarten is not a significant predictor of arithmetic in 1st grade when age and intelligence are controlled for. It explains an additional 1,7% of the variance in arithmetic. After adding number sense, age no longer predicts arithmetic, but intelligence still do. Model 2 itself is significant (i.e., $F(1, 156) = 8.28$, $p < .001$).

In Model 3, early numeracy was added to see if it would mediate the relation between number sense and arithmetic. The results showed that early numeracy explains an additional 10.6% of variance in arithmetic. After adding early numeracy to the model, intelligence also no longer predicts arithmetic. Early numeracy is now the only significant Kindergarten predictor of arithmetic in 1st grade. Model 3 itself is significant (i.e., $F(1, 155) = 8.28$, $p < .001$). The beta values of the final model show that an increase of the standard deviation in early numeracy will result in an increase of .36 standard deviation in arithmetic, with a significant level of $p < .001$.

Table 6.

Explained Variance per Model

Model	R^2	Adjusted R^2	$\Delta R2$	*F*	*df1*	*df2*	*p*
1	,095	0,84	-	8,279	2	157	,001
2	,112	0,95	,017	2,919	1	156	,090
3	,218	,198	,106	20,947	1	155	,000

Model 1: Age, Raven

Model 2: Age, Raven, Dot Comparison (K)

Model 3: Age, Raven, Dot Comparison (K), Early Numeracy

4.6 Summary of the analysis and results

The correlation analysis showed that number sense had a weak but significant correlation with arithmetic. Early numeracy had the highest correlation with arithmetic, which was moderate in size. Non-verbal intelligence and age both had a weak but significant correlation with arithmetic. Vocabulary had a weak a non-significant correlation. Therefore, the independent variables number sense, early numeracy, intelligence and age were included in the next step -the regression analysis - to see to what extent they could predict the dependent variable, arithmetic.

In the final step of the analysis, a hierarchical multiple regression analysis was carried out. Although a significant correlation between number sense and arithmetic had been established, it was discovered that when controlling for intelligence and age, number sense did not have a unique effect on arithmetic beyond intelligence and age. Thus, the results of this analysis indicate that number sense could only to a small degree explain variation in arithmetic skills. The third and last step of the model shows that early numeracy was a significant predictor of arithmetic, and it explained 10.6% of the variation of arithmetic abilities in Norwegian 6-year-olds, after it is controlled for age, non-verbal intelligence and number sense in kindergarten. Therefore, it can be concluded that early numeracy at age 5 is the most important predictor for arithmetic at age 6. The results of the analysis will be discussed further in the next part in light of the other empirical findings, validity theory, and educational implications.

5 Discussion

In recent years, a lot of research has been carried out into the role of number sense in the development of arithmetic and mathematics. The findings are conflicting, and therefore the purpose of this study was to look further into the role of number sense in predicting arithmetic abilities. The aim of this study was to investigate to what extent number sense at age 5 can predict arithmetic abilities of Norwegian speaking monolingual 6-year-olds. Because of the conflicting findings in the literature, early numeracy was controlled for to assess the possible influence of that variable as well. Many studies have reported abilities within early numeracy that are significant and unique predictors of arithmetic. Non-verbal intelligence was also controlled for, as it is a well establish predictor within the area of mathematics. Vocabulary was also assessed, to establish that there were no other reasons for presence or absence of a statistical relation between precursors of mathematical ability, i.e., number sense and early numeracy, and the outcome variable arithmetic. Therefore, a longitudinal study was conducted, to assess what kind of early abilities play a predictive role in the acquisition of arithmetic in 1st grade.

A review of the literature has shown that the development of arithmetic abilities is a cumulative process, where one skill sets the foundation for the next. The development starts with preverbal abilities (Geary, 2000). These are the abilities known as number sense (Dehaene, 1992), including subitizing and approximate number sense, which children acquire before they learn to speak (i.e., preverbal stage). Subitizing is the ability to quantify small sets of objects (Anuio & Räsänen, 2015). Approximate number sense is the ability to estimate which amount is greatest, in larger sets of objects without having to count (Libertus et al., 2011). It is believed that this preverbal and innate ability to perceive quantities is necessary in order to acquire a more complex numerical understanding (Geary, 2000). These abilities lay the foundation for early numeracy, which is the increasing ability to comprehend mathematical relationships. In the beginning, early numeracy is about learning and understanding numerals and numeric frequencies, e.g., learning to count by using number words (Desoete, 2015). This then goes on to linking number words to Arabic numerals. Altogether, this eventually becomes the ability to combine number words with the Arabic numerals, including an understanding of the number amounts (e.g., three = 3 = ***). From early numeracy the child develops an understanding of arithmetic. Early on, children can use counting strategies like adding (1 plus 3 = 4) by using fingers or objects. These skills are then transferred into the ability to calculate with Arabic numerals and mathematical symbols such as +, -, and = (Siegerl & Braithwaite, 2004). Thus both abilities, number sense by estimating quantity and early numeracy as the ability to use numbers for quantity, have important roles in the development of early arithmetic.

The extent to which this representation of mathematical development is correct and can be generalized across generations and cultures is uncertain, although some similarities have been found (Geary, 2000). For instance, some traits seem so develop similarly independent of time and place. These are abilities, such as numerosity, ordinality, counting, and simple arithmetic, that are found in preverbal infants and young children (Geary, 2000). The development of mathematical abilities thereafter then varies and is to some extent culture bound. These are the abilities that together constitute early numeracy. In addition, theories suggest that a failure in the understanding of quantity, verbal abilities and/or visual numbers can lead to different subtypes of developmental dyscalculia (DD). Since these are abilities that develop at different stages, DD can occur at different points during development. There is also considerable disagreement in the literature and in the empirical findings if it is the abilities within number sense or the abilities within early numeracy, such as counting, number identification, seriation and classification that are important for the development of arithmetic abilities. In recent years, some research has shown how number sense relates to arithmetic skills, and how children with weak abilities in estimating quantities also develop weak arithmetic skills. However, other studies indicate that there are more specific abilities within early numeracy that are crucial to the development of arithmetic skills. Others again present a proposal on how number sense and early numeracy together play an important role in the acquisition of arithmetic skills, and that number sense at an early age is important for the development of early numeracy and is therefore important for the development of arithmetic abilities.

The results of the current study are discussed next, together with the empirical findings. The study’s validity and reliability will be discussed in relation to the possibility to generalize the findings, and the limitations of the findings. Also, there are some methodological limitations that need to be addressed. There may be relationships found here that can be explained by other variables that have not been accounted for, such as working memory, processing speed, motivation and family socio-economic background, also known as general-domestic skills. In addition, a discussion will be given on the relevance of the results for the educational field, before a brief summary is given and then a look of the way forward.

5.1 The results considering theory and empirical findings

The findings of this study show that number sense in Norwegian 5-year-olds in Kindergarten has a weak to moderate, but significant, relation with arithmetic skills one year later in 1st grade. This is in line with previous studies that have identified the same relation between number sense and arithmetic abilities (Bonny & Lourenco, 2013). Likewise, it was detected that non-verbal intelligence and age were weakly to moderately related to arithmetic. The strongest relation was found between the early numeracy in Kindergarten and arithmetic abilities in 1st grade, which was moderate. Additional analyses showed that the relation between number sense and arithmetic is not predictive in nature. Number sense at age 5 does not explain a significant amount of variance in arithmetic abilities one year later, when controlling for non-verbal intelligence and age. Of the precursor variables from Kindergarten, only early numeracy was found to predict arithmetic abilities in 1st grade.

The results of previous studies that have researched number sense's role in arithmetic abilities vary. Some studies have found that number sense can predict later arithmetic skills (Desoete; et al., 2013; Libertus et al., 2013; Mazzocco et al., 2011; Toll et al.,2015), while others have found correlations between number sense and arithmetic, but that number sense is not predictive for arithmetic (Göbel et al., 2014; Sasanguie et al., 2013). One way in which these different results could be explained is by means of the properties of the tasks. Various methods have been used to test the properties of number sense and, more precisely, approximate number sense (ANS). ANS can be tested by estimating the amount that is the greatest of two quantities and the task can be carried out on a computer or on paper. Within this kind of task, either accuracy, ratio (Weber faction) and/or response time are assessed. In other versions, a similar test is carried out by comparing a quantity with two other quantities, where the participant must find the quantity that is equal to or closest to the first quantity. This is also done on a computer or paper. In the current study, number sense is tested with the part ability approximate number sense. This is done with a subtest from TOBANS (Brigestocke et al., 2016), a dot comparison test where the child gets 30 seconds to estimate the greater amount of two quantities on paper. That means, that number sense was tested by accuracy of ANS acuity.

The results mean that the preverbal abilities measured by ANS in this study are related to arithmetic, but it does not have a predictive role in the development of arithmetic skills.

Rather, it is early numeracy that plays a decisive part. Such a discovery is not surprising as several previous studies have found that early numeracy or part abilities within the area are unique predictors or found as important predictors along with other abilities such as number sense. Arithmetic abilities are in this study tested with fluency addition and fluency subtraction, with a subtest from TOBANS (Brigestocke et al., 2016). Here, the result of the study will be discussed against the empirical findings, while taking into account the specific properties of the tasks used in the current study, and, like the theory part, it will be divided according to the applied research design.

5.1.1 Cross-sectional studies

Libertus and colleagues (2011) found that ANS is significantly related to children's early mathematical skills at age 3 to 5 (M_{age} = 4.20), while controlling for age and vocabulary. They conducted a dot comparison task on computers using accuracy, Weber fraction and response time to measure ANS. Mathematical skill was measured using TEMA-3 (Test of Early, Mathematics Ability) consisting of verbal counting with one-to-one correspondence, number comparison, numerical literacy, retrieving of number facts, arithmetic skills and number concepts. This study deviates from the current study in multiple areas. For instance, they measure multiple mathematical abilities, where one of the abilities is arithmetic, similar to arithmetic fluency measure in the current study. Age wise, the current study deviates from Libertus and colleagues (2011) where the children in the current study have a mean age of 5.49, which is over a year older, than the children in the other study at the first measuring point. Libertus et al. (2011) found that ANS has a significant correlation with mathematics using all three units of measurement. Furthermore, they found after running a regression analysis that ANS accuracy, Weber fraction and response time all contributed to a significant relation between math abilities and ANS acuity. There are two main deviating factors which can contribute to explaining why Libertus et al. (2011)'s results are different from those of the current study; age difference of the participants and the outcome measure of ANS acuity. Firstly, in order to measure math abilities, they used, among other, the abilities of counting with one-to-one correspondence, reading Arabic numerals, as well as addition and subtraction tasks. In my study math abilities are divided into two skills, early numeracy consisting of 'count on' and reading of Arabic numerals in kindergarten, and arithmetic consisting of addition and subtraction fluency tasks in 1st grade. Therefore, a significant relation between ANS acuity and math abilities in their study, might be a result of the us of a range of mathematical abilities, including abilities within early numeracy and arithmetic. Since the development of mathematics skill is cumulative (i.e., number sense → early numeracy → arithmetic), measuring math abilities within early numeracy concepts can give a significant relation between ANS and math abilities. As ANS and early numeracy are closer in the developmental stage than ANS and arithmetic measured in the current study. Furthermore, it is of interest that the children in their study are younger than those in this current study, as it is possible that measurements of numbers sense and approximate number sense are more relevant for younger children, before they acquire early numeracy to a greater extent.

Bonny and Lourenco (2013) wanted to see if younger children's ANS abilities correlated to a greater extent with their mathematical abilities than for slightly older children. Therefore, they tested children ranging in age from 3 to 5 years and compared their ANS skills with their mathematical abilities, by dividing them into 3 groups according to their age, i.e. 3, 4, or 5 years. They tested ANS with a dot-comparison task measuring accuracy, as in the current study. For math abilities, they used TEMA-3, testing verbal counting, cardinality and numerical arithmetic abilities. They found that children age 3 and 4 showed a significant relation between their ANS acuity and mathematical abilities, such as counting, Arabic numerals and arithmetic, while there was no significant relationship between ANS and math abilities in children at age 5. Furthermore, they found that the better the children were in mathematics, the weaker the relation was with ANS. These results can then be in line with both Libertus and colleagues' (2011) results and those in the current study. Thus, the younger children in Bonny and Libertus (2013) show that there is a significant relationship between ANS and mathematics, in line with Libertus and colleagues' (2011) findings, in addition to finding no significant relationship between ANS and mathematical skills in children aged 5, consistent with the outcome of the current study.

Fuhs and McNeil (2013) did not find a significant relation between ANS acuity and mathematical abilities in children age 3 to 6 (M_{age} = 4.58). They used an object comparison task, measuring accuracy and Weber fraction of ANS. Math abilities were tested with TEMA-3 consisting of testing relative magnitude, counting skills, number and numerical calculation skills, number facts and base-10 concepts. They wanted, in line with the other studies, to see if ANS predicts mathematical skills, but focused specifically on children from low socio-economic backgrounds. They found, in line with the current study, that ANS acuity has a weak, but significant correlation with mathematical skills. However, though when only looking on ANS accuracy, the correlation was weak and non-significant. Furthermore, controlling for receptive vocabulary and inhibitory control caused the significant relation between ANS acuity and math abilities to disappear. Based on the theory they used, it is possible that children with low-socio-economic status are exposed to less mathematics than other children, and therefore develop the various skills within number sense, early numeracy and arithmetic at a slower rate, because all three are skills that increases with age and by exposure, like during education.

5.1.2 Longitudinal studies

As in the cross-sectional studies, the results of longitudinal studies vary. Mazzocco and colleagues (2011) assessed, in line with the current study, if ANS precision in 3- to 4-year-olds, before starting school can predict school mathematics two years later. ANS acuity was measured with an object-comparison task using Weber fraction. Math abilities were tested with TEMA-3, consisting of abilities such as counting, Arabic number identification, and addition and subtraction tasks. They found that ANS in 3- and 4-year-olds is a significant predictor for mathematical knowledge two years later, when controlling for full-scale IQ. This result is in line with the results of Libertus et al. (2011). Both studies did not test for early numeracy and they used the same abilities to measure mathematical abilities. Their results conflict with the result of the current study. As pointed out, Mazzocco et al. (2011) used Weber fraction, in contrast to the current study that used accuracy. Furthermore, they measured math abilities, combining abilities that in the current study are divided into abilities within early numeracy and arithmetic. The used measurement unit for ANS i.e., Weber fraction, and the outcome variables, makes it difficult to compare their result with the current study. In addition, the result from their study is difficult to generalize, in that the sample size consists of only 17 children. Yet it is worth mentioning in the bigger picture, because it shows results coinciding with those of Libertus et al. (2011) and those of the longitudinal study to Libertus and colleagues (2013, described below).

Libertus and colleagues (2013) aimed to investigate whether the response time and accuracy for ANS of 204 preschoolers (M_{age}=4.20 years) could predict mathematical skill six months later. ANS was measured with a dot-comparison task conducted on a computer, and math abilities were measured twice with a 6-month gap. For testing math, they used TEMA-3 that consists of verbal counting with one-to-one correspondence, reading Arabic numerals, number comparison, fact retrieval, mental calculation, and symbolic arithmetic tasks. They found that both response time and accuracy for ANS were significant and unique predictors for math abilities 6-month later, even when they controlled for early mathematical abilities from the first measuring point, such as counting, Arabic numerals, and arithmetic skills. This result is in contrast with the current study, even though, they have used similar measurements. As in the current study, they measured ANS accuracy, though on a computer, and they used math abilities similar to the abilities within early numeracy, i.e. counting and reading of Arabic numerals. Libertus et al.'s (2013) study contrasts with the current study regarding at the age group, where their children had a mean age of 4.20, the children in the current study had a mean age of 5.49, resulting in a whole year difference. Furthermore, they have not controlled for age and non-verbal intelligence, which was controlled for in the current study. Controlling for this in the current study resulted in that ANS did not contributing a significant amount of explained variance in arithmetic abilities in the current study. If they had tested for age and non-verbal intelligence their result could have been different. In addition, their outcome variable contains several math abilities, while the current study only measured arithmetic fluency.

The results of Mazzocco et al. (2011) and Libertus et al. (2013) are very similar, for testing ANS they used different measurement units, where Libertus et al. (2013) used accuracy, like in the current study, and Mazzocco et al. (2011) used Weber fraction. They used the same standardized test TEMA-3 for testing math abilities. They have also tested children who are about the same age, from 3- to- 4 in Mazzocco et al. (2011)'s study and Libertus et al. (2013) had a mean age of 4.20, in contrast to the current study that had a sample with a mean age of 5.49. Their results do not match the results from the current study. The question about age is a crucial factor here as their children are about one year or more younger at the first measurement point than in the current study. Both studies used TEMA-3 for the second measuring point. Here, they included abilities that are near number sense in the mathematical developmental process. This can possibly make the relation between ANS and math abilities greater, than if they had only measured abilities acquired at school age, such as arithmetic. The question about the importance of measurement unit of ANS slightly decreases when comparing these studies. All have found a significant relationship between ANS and mathematical skills regardless of whether they used accuracy, response time or Weber fraction. Especially, the result of Libertus et al. (2013) builds on this idea as they found that both RT and accuracy were individual significant predictors for mathematical abilities.

Desoete and colleagues (2012) examined whether individual non-symbolic and symbolic comparison skills in children at age 5 to 6 could predict arithmetic skills two years later. In addition, they divided the children into three groups, typical achievement (TA), low-achievers (LA <25th percentile) and mathematical difficulties (MD <10th percentile). Non-symbolic comparison abilities were tested with a dot-comparison task, similar to the one used to measure number sense in the current study, measuring accuracy. Symbolic comparison skills were tested with different tasks, e.g., by judging if a word was a number word, and comparing two number words and repeating the larger one. Arithmetic skills were measured in the first and second grade, with addition and subtraction tasks on paper and orally. For the whole sample, they found that non-symbolic comparison abilities predicted arithmetic skills one year later, and fact retrieval two years later, in contrast to my result. They found that weak abilities in ANS and Arabic numerals predict later mathematical skills more strongly in the MD and LA group, while non-symbolic comparison abilities do not predict arithmetic abilities in typical achievers. This is in line with the current study. Similar tests and measurement units have been used for testing both number sense and arithmetic abilities, and for the control variable non-verbal intelligence. In addition, they had a sample with a similar age range (5- to 6-year-old) as the current study.

Göbel and colleagues (2014) aimed to identify what abilities at age 6 (M_{age} = 6,30) have a longitudinal predictive effect on arithmetic skills eleven months later. ANS was measured with a non-symbolic comparison tasks, where the child compared two amounts within squares of the same size. In the other non-symbolic comparison task, the child had to compare two quantities placed on the same size surface area. Both non-symbolic comparison tasks measured ANS accuracy. For measurement of abilities within early numeracy, they used a number identification task, where the child heard a one-, two- or three-digit number word and had to identify the corresponding Arabic numeral. Furthermore, they used Raven to measure non-verbal intelligence and the British Picture Vocabulary Scale (BPVS III) to measure vocabulary knowledge. For measuring arithmetic abilities, they used addition, subtraction and multiplications problems, which increased in difficulty level from measurement timepoint 1 to timepoint 2. To start, the study by Göbel and colleague (2014) is very similar to the current study. They use the same concepts for the longitudinal predictors of arithmetic, even though it does vary a bit. They used a similar task for measuring ANS and used the same measurement unit, accuracy. For early numeracy they used number identification, while the current study used number reading and count on. Number identification requires similar abilities to those required in number reading. For arithmetic they used calculations with addition and subtraction, but without a time limit, and did not measure arithmetic fluency, as was done in the current study. They used, however, exactly the same measurement tools for non-verbal intelligence and vocabulary knowledge. The difference in age range from their study to this study is only half a year (M_{age} = 5.49 and M_{age} = 6.30). In all, Göbel et al. (2014)’s study is very similar to the current study. They found that the ANS at point one and arithmetic abilities at time point two were strongly related. This is similar to the findings of the current study, though that correlation was weaker. Furthermore, they found in line with the study, that early numeracy was stronger related to arithmetic abilities than ANS. They found, like the current study, that ANS and skills within early numeracy, such as number comparison and number identification have a strong relation with arithmetic skills. In addition, they found that only number identification was a unique predictor of arithmetic skills.

Toll and colleagues (2015) examined the relation between growth in non-symbolic and symbolic comparison skills, as well as whether or not non-symbolic and symbolic skills of kindergartners (M_{age}=4.60) predicted basic mathematical skills up to 2 ½ years later in 1st grade. Non-symbolic comparison abilities were measured with a dot-comparison task to measure ANS accuracy, similar to the one used to measure ANS accuracy in the current study. Mathematical abilities were tested by testing numbers, number relations and simple addition and subtraction tasks verbalized by the teacher. Arithmetic abilities were tested with an addition- and subtraction fluency task, similar to the one in the current study. First, they found that ANS at age 5.60 years (mid-year, second timepoint in Kindergarten) has a weak, but significant relation with arithmetic fluency skills in first grade, concurring with the result in the current study. In all, they found that ANS from Kindergarten predicts arithmetic fluency in 1st grade. However, symbolic comparison skills were the main predictor of the broader math abilities. The result that ANS predicts arithmetic fluency is in contrast to the result in the current study, which found that ANS does not explain variance in arithmetic fluency.

The longitudinal studies of Desoete and colleagues (2011) and Göbel and colleagues (2014) had results fairly similar to those of the current study by using similar concepts to the operationalized terms for number sense, early numeracy and arithmetic. In addition, Göbel et al. (2011)’s study used the same covariates as the current study, i.e. non-verbal intelligence and vocabulary. Both studies found, concurrent with this study, that abilities within early numeracy i.e., symbolic verbal number words, number identification, number reading and counting on, all have stronger predictive value for arithmetic abilities, than other control variables, i.e., number sense, non-verbal intelligence and age. These findings make the results of the current study more valid. However, Toll and colleagues (2015) discovered, while using the same concepts as in the current study, that ANS accuracy from Kindergarten does predict arithmetic fluency abilities at the end of first grade. This is opposite to the findings in the current study. On the other hand, Toll et al. (2015) also found that symbolic-comparison skills were the main predictor for broader math abilities. Looking at the findings of these three studies and the current study, the results coincide, and deviation seem to derive from age differences between the samples. Toll et al. (2015) started measuring children that were a year younger (age 4) than the other studies. Also, by looking at Desoete et al. (2011), it can seem that it is the stage the children are at in their mathematical development that is decisive in whether ANS has a predictive value or not.

5.1.3 Experimental studies

The cross-sectional and longitudinal studies display varying results. Some indicate that number sense plays an important role in the development of arithmetic and mathematical skills, while others find that the role of number sense disappears when they control for other factors, especially early numeracy. A recurring factor in the studies is the age difference of the participant groups. The experimental studies discussed in the next part, have a more variable age range. All of the studies investigate the role of ANS in arithmetic skill, with the exception of one that investigated what makes ANS develop and become more precise.

The latter is a study by Piazza and colleagues (2013) that examined whether it was age or education, or both, that was a decisive factor in the development of ANS. The sample had an age range from 4 to 63. They measured the participants ANS with a dot-comparison task conducted on computers, and used Weber fraction as the unit measurement. They found that the level of education, especially in mathematics, were a significant factor for the development of ANS. This suggests that ANS is something that can be trained by what Geary (2000) categorizes as secondary abilities. Secondary abilities are abilities acquired at school, such as symbolic arithmetic. Given the previous studies we have looked at, this raises the question of how important the age of the children is in relation to how much mathematical knowledge they have been exposed to. Several of the studies have shown a relationship between ANS and arithmetic skills if the children are assessed early enough in development. This may indicate that it is not the age itself but their acquired knowledge that is crucial. This can also explain why the research shows a link between ANS and arithmetic skills in children who are LA or have MD/DD. In relation to the current study, the findings can be a result of the fact that the children have already acquired the components within early numeracy, making it more important than ANS abilities.

Hyde and colleagues (2014) did two experiments on first graders (M_{age} = 6.90) to see if activating the ANS could enhance children's performance of symbolic arithmetic. To train ANS the children received one of four training tasks; non-symbolic addition, line length addition, non-symbolic numerical comparison and brightness comparison, all conducted on a computer. They used ratio, response time and accuracy as unit measurements for ANS. First, they found that the pupils within the non-symbolic numerical addition and non-symbolic numerical comparison groups, were faster to complete arithmetic problems after having received training. They found further that the training of ANS only enhance arithmetic abilities, by controlling for vocabulary. They concluded that training of ANS improves pupils' ability to solve arithmetic tasks, indicating that there is an overlap between symbolic arithmetic and ANS. This result is in contrast to the result of the current study. If ANS does not have any unique predictive value within this age group, then enhancing ANS should not result in a significant improvement in arithmetic performance.

Obersteiner and colleagues (2013) had a similar result from their experiment, where they checked if training of the exact (subitizing) or the approximate number sense would enhance arithmetic performance. They trained 204 first graders' (M_{age} = 6.91) number sense by dividing them into different groups, to train exact number sense, approximate number sense, both exact and approximate number sense and one control group receiving language training. All training was conducted on computer, using dot-comparison tasks. Number sense was measured using accuracy and response time as measurement units. Obersteiner and colleagues (2013) found that training of both exact or approximate number sense improved the math performance of first graders. The result of this study is in agreement with the study of Hyde et al. (2014) and is therefore also conflicting with the result of the current study.

Lastly there is the experiment of Wang and colleagues (2016) that tested if ANS can be both enhanced and disrupted, and if so, if it affected children's symbolic math performance at age 5 (M_{age} = 5.30). They trained the accuracy of ANS acuity with a dot-comparison test conducted on a computer. First, they found that ANS acuity can be altered, meaning that it can be both enhanced and disrupted. Furthermore, they found that the those who had enhanced their ANS acuity did better in the math test and those who had their ANS disrupted did poorer in their math test.

The three last studies by Hyde and colleagues (2014), Oberteiner and colleagues (2013) and Wang and colleagues (2016) all have results that concur with each other, but are not in agreement with the current study. As stated earlier, if ANS does not have unique predictive value, as the current study finds, then enhancing ANS should not result in a significant enhancement of arithmetic performance. However, there is an important factor that discerns their result from the current study, and that is their chosen control variable. While they controlled for vocabulary, to make sure that enhancing ANS only affected the mathematical domain, they did not control for abilities within early numeracy, as controlled for in the current study. Further experimental studies on whether enhancing ANS effects arithmetic abilities, could include a control group that receives training on abilities within early numeracy to see if the group that receives early numeracy training gets a better, equal or smaller enhancement of arithmetic abilities, compared to the ANS group.

5.1.4 Summary/Conclusion

The experimental studies indicate that there is a causal link between ANS acuity and mathematical abilities. Along with the cross-sectional studies, longitudinal studies and the current study, it seems that whether ANS predicts arithmetic skills, or not, depends on what stage the child is at in their development and not necessarily their age, per-se. This means that when the child reaches a certain level of mathematical knowledge, there are other abilities within early numeracy that take over from number sense and therefore become more crucial to the development of arithmetic skills. This hypothesis could explain the divided findings within the research field examining ANS's role in mathematical development.

5.2 Validity and reliability of the study

5.2.1 Construct validity

In education research, the concepts that are investigated cannot be directly observed, and therefore need to be operationalized. The question of concept validity is then about the extent to which the theoretical concept matches the operationalization that is used. An assessment of concept validity is therefore very relevant to this study (Shadish et al., 2002). In this study, the concept validity is linked to the variables: number sense, early numeracy, arithmetic, non-verbal intelligence and vocabulary.

Number sense in this study is operationalized as approximate number sense (ANS). ANS is a part of number sense, leaving out exact number sense (Dehaene, 1992). A dot-comparison task from TOBANS (Brigestocke et al., 2016) was used to measure ANS. Although the dot comparison test has been translated into Norwegian, this is not considered a threat of limited validity, as the test is not considered to be vulnerable to translation. The outcome used for ANS is in this case accuracy, which can be a weakness since ANS can be tested with accuracy, Weber fraction and response time. Despite this, a closer look at the empirical findings has shown that the different outcome measurements of ANS tasks seem to produce similar results. Because of this, the applied measuring tool for ANS is considered good, and therefore does not threaten concept validity.

To test early numeracy, two tests were combined, reading numbers and counting on. Early numeracy is a collective concept of multiple skills and, in this case, not all early numeracy skills were included. Among other, we did not test abilities such as one-to-one correspondence and cardinality. However, the tests that were used are considered to be sufficient to test early numeracy covering both the auditory verbal code and the visual Arabic code (Dehaene, 1992).

Arithmetic skills were operationalized with two tests, addition fluency and subtraction fluency from TOBANS (Brigestocke et a., 2016). These tests come from the same test battery as dot comparison, and are therefore also translated, but this is not considered a threat here either. Although this test has a time limit, it is still considered good enough to test arithmetic skills, as automatic retrieval of arithmetic facts is a part of basic arithmetic abilities.

For the covariates, non-verbal intelligence was used as a control variable and is operationalized with the Raven (Raven, 1998). Raven is translated into Norwegian. Since it is a test that does not require verbal skills, it is considered to have good concept validity. BPVS was used to test vocabulary and is also considered to have good validity (Dunn et al., 1997).

Based on this, the conceptual validity of the study is considered good, as the tested tests used to operationalize the constructs are recognized and similar to the tests that have been used in other studies investigating the same concepts.

5.2.2 Statistical validity

Statistical validity is about the statistical strength between the dependent and the independent variables. The statistical strength determines how reliable the conclusions are. Violation of assumptions in the analysis may pose a threat to the statistical validity. In this study, the prerequisites for being able to carry out correlation analysis, as well as regression analysis, have been carefully reviewed, and no deviation from the assumptions was discovered. However, it cannot be ruled out that external variables may have affected the results of the analysis to some extent.

Low test or measurement reliability is also a threat to statistical validity. To test this, the reliabilities of the instruments used in this study were evaluated using Cronbach's alpha. Reported values of Cronbach's alpha were all based on the study sample. All the tests reported a good reliability above .70. This means that the internal consistency of the test is considered to be good (De Vaus, 2014) and does not pose a threat to the statistical validity of this study.

Another threat is systematic measurement errors. Systematic measurement errors will result in wrong outcomes, and then constitute a reliability threat. This can come from a test assistant who administers a test in an incorrect way, or from errors in the test itself. In the project to which this study is connected, all administrators were master's degree students who collected the data after careful training. In addition, every test situation was recorded on tape. The audio tapes have been reviewed by other employees in the NumLit project to ensure that no systematic errors have been made; and in cases where these were detected, test scores or whole participants have been removed. Therefore, systematic measurement errors are not considered a threat in this study.

Overall, the statistical validity in this study is considered to be good.

5.2.3 Internal validity

Internal validity refers to the degree to which causal conclusions can be drawn. It addresses the direction problem, i.e., if the dependent variable is an effect of the independent variable or vice versa. In this study it is not possible to draw causal conclusions because of the choice of method. In this study, a relationship between the variable arithmetic and the variables early numeracy, number sense, age and non-verbal intelligence was detected. Since secure conclusions about causality cannot be drawn with the chosen research design, alternative interpretations of validity must address the direction problem and the threats of third variables (Shadish et al., 2002).

This study aimed to look at how ANS can predict arithmetic skills after one year in Norwegian 6-year-olds. To do this, other variables that have been proven to be important predictors from previous research were controlled, in this case, early numeracy, non-verbal intelligence and age. It was found that number sense, early numeracy, non-verbal intelligence and age were significantly related to arithmetic. In addition, age and non-verbal intelligence were found to be significant predictors of arithmetic skills in 1st grade, and that number sense did not explain a significant amount of variance in arithmetic abilities. Additionally, it was discovered that early numeracy is a significant predictor after taking the variables named above into account. It cannot be said with certainty that pupils' arithmetic skills are the result of their early numeracy, as the causal conclusion cannot be drawn. With that in mind, it is nevertheless entirely possible that early numeracy is one of the foundational skills needed for successful mathematical development, considering our current knowledge about mathematical development. This is, because we know that math abilities develop gradually over years, and early numeracy skills are learned before symbolic math and are considered important precursors for the acquisition of arithmetic knowledge.

Possible influence of third variables is another and significant threat to the internal validity of this study. The found statistical relation may come from a hidden third variable that can affect arithmetic skills. In this study, it was therefore controlled for that a possible statistical relation between number sense and arithmetic did not come from other variables such as early numeracy, non-verbal intelligence and age. The result was that number sense was not a unique significant predictor when controlling for the other variables, but rather that early numeracy was a unique predictor. Despite this, there may still be other hidden variables that are crucial. Previous research has found that, among others, working memory, processing speed, socio-economic background and motivation are also important factors in the early development of mathematical ability. In this study, these factors have not been accounted for. However, the control variables non-verbal intelligence and age are considered as good and appropriate when investigating the role of cognitive factors, as both variables have been found to be important factors in previous studies. Including these covariates, has helped to reduce the threat of third variables to internal validity.

Although this study cannot directly test causal relation, internal validity is considered satisfactory with the support of previous research and empirical findings, as well as the control variables used.

5.2.4 External validity

External validity concerns the extent to which the outcome of the study can be generalized and the extent to which the sample results are representative for the population. This means that in order to have good external validity, it depends on relationships that can be generalized to a wider context or other contexts (Shadish et al., 2002).

A random sample is important to have good external validity. The sample in this study is randomized within the selected municipalities. This can be said when all the kindergartens that fulfilled the inclusion criteria had the opportunity to participate. However, it must be taken into account that there may be selection bias that can lead to a skewed selection. This may be due to the fact that parents who consented to their children participate in the study may represented a special group, and that the parents' background may have affected individual conditions for the children in the selection. Also, in the ethical part (see Part 3.6) of this thesis, we posed the question of whether a child may have been negatively affected by external factors before the test session; could it affect the pupil’s performance? In this current example, some of the pupils were addressed in a derogative way. If the experienced example is a repeating event, across schools and municipalities, it can result in a distorted image of the result to a pupil group, giving a systematic selection. These are of course assumptions that cannot be verified. Furthermore, the sample in this study is supposed to represent the normal variation within the target group, as no systematic selection of participants was carried out. This can help to ensure that there is no large individual homogeneity in the sample, and therefore the external validity is considered satisfactory at this point. The size of the sample in this study further strengthens the external validity.

Before carrying out the analysis, assumptions were checked. As none of the assumptions were violated there is no large threat to external validity. Furthermore, low statistical validity may impair external validity if the found statistical relationships are not significant. In this study, statistical relationships have been found to be significant, and are not considered a threat.

With this, the external validity is considered good, and therefore the results of this study can be generalized with certain reservations.

5.3 Educational implications

One of the school's most important tasks is to understand which factors affect the development of pupils' mathematical skills. The theory used in this study has shown that these can be external factors, such as teaching, home environment, or more closely related factors such as the pupil's abilities and capacity. In a broader view, this means the pupil's expected ability to acquire mathematical skills. This includes, among others, secondary and domain-general abilities (Passolunghi & Lanfrachi, 2012). Additionally, there are specific abilities that have been shown to influence mathematical skills, i.e., the domain-specific abilities. While it is important to understand how secondary relationships and domain-general skills affect arithmetic skills, the focus of this study is on the capabilities within the domain-specific ability area, and the importance of the various capabilities regarding the acquisition of arithmetic. Domain-specific abilities are, according to the triple code model (Dehaene, 1992) abilities that affect numerical processing and that are divided in three modules. 1) the analog magnitude code represents the ability to understand approximate amounts, quantity within counting and numerals, in this study represented as number sense. 2) the auditory verbal code addresses numerals, counting strategies and fact retrieval, represented here as the ability to 'count on' as a part of early numeracy and 3) the visual Arabic code deals with the ability to manipulate and represent numbers in Arabic format, which in this study is represented as the ability to read Arabic numerals. A failure within any of these three modules could cause difficulties in the acquisition of arithmetic skills, causing developmental dyscalculia (Butterworth et al., 2011).

Based on this theoretical framework, approximate number sense plays an important role in the mathematical development of younger children (Aunio & Räsänen, 2015: Dehaene, 1992; Geary, 2000). According to the findings in this study, number sense measured as ANS has a relation to arithmetic abilities, though it is not predictive in nature. A cause of this can be that the children in the sample have reached a certain stage in their mathematical development, i.e., early numeracy, that makes ANS less important. Therefore, interpreting this study and the empirical findings, it can be hypothesized that a preverbal failure, that is, a failure within the analogue magnitude code, in the form of (approximate) number sense, must be detected at an earlier stage in the child's development to prevent detrimental effects on arithmetic and math abilities. According to this, if a child is lagging behind, and is at a developmental stage that is expected for children at age 3 or 4, it will be justified to focus on enhancing the child's ANS abilities.

The findings of this study show that it is early numeracy that is the decisive cognitive factor in arithmetic abilities at the end of Kindergarten. That is, early numeracy was a unique significant predictor of arithmetic at age 6, when controlling for number sense, non-verbal intelligence and age. This is consistent with empirical findings (Göbel et al., 2014; Toll, et al., 2015). Early numeracy is a set of different skills and in this study represented as the ability to ‘count on’ and read Arabic numerals. The ability to “count on” requires the child to know the numbers words and the number sequence, by answering what number comes next in line, i.e., "two, three, four and then comes…?". The counting skills that are required here are important/basic skills that children should develop early, through different stages to acquire counting strategies (Aunio & Niemivirta, 2010). In the ‘count on’ test, the child uses its auditory verbal code to count, but is not required to grasp one-to-one correspondence or have an understanding of cardinality, which are also key abilities within counting, and hence early numeracy. The second test measuring early numeracy in this study was reading Arabic numerals. This tests the child's ability to recognize Arabic numerals and to say the numbers correctly. This also tests the child's knowledge of the number system, such as knowing that forty-two is written as “42” and not “402”. The abilities within early numeracy are, according to the findings in this study, important abilities for the future development of arithmetic abilities. Therefore, focusing on strengthening early numeracy towards the end of Kindergarten is essential to ensure future successful arithmetic development.

Early numeracy in this study does not include every skill, but it does cover ‘number reading’ and ‘count on’, testing skills within the auditory verbal code in terms of counting and the visual Arabic code in the form of number reading. What is missing is a measurement of the three codes combined as it is important to ensure an understanding of the relationship between the number word, the Arabic numeral and the amount, i.e. “three = 3 = ***”. This could have been tested with counting objects with one-to-one correspondent and cardinality to get a better picture of the child's early numeracy. With such a comprehensive overview, employees and parents can get an insight into the child's understanding of these mathematical components.

In Norway, there is an ongoing debate about children’s right to free play and the need to prepare for academic requirements. Free play has a big role in Norwegian culture and is considered to be the child’s main right. Therefore, the increasing focus on academic and formal education in kindergarten is giving rise to a lot of uncomfortable feelings. However, stricter demands as to what a child in kindergarten should learn are coming from the Ministry of Education (2017). The Norwegian framework plan (Ministry of Education and Research, 2017) for kindergarten, does not have a specific part for mathematics, but does include mathematical principles as a subpart, and the wording can be freely interpreted by every kindergarten. It states that the kindergarten should facilitate the child in discovering and exploring mathematics in everyday life, with emphasis on quantities, numbers and counting.

The focus of the framework plan is then in line with the findings in this study; if the kindergarten teaches the child about quantity, numbers and counting, and the connection between the different representations (i.e., ** = two = 2), they provide the child with a good framework for later learning. However, this requires some effort and willingness by the kindergarten and employees who have knowledge of mathematical development. An article by Storvik (2019) states that children from high-quality kindergartens are one-year a-head in the development of basic abilities (i.e., math, reading, writing, oral and IT) and social competence, compared to children from low-quality kindergartens. If the kindergarten maps any deviations in development of quantity, number or Arabic numeral skills, they can detect and possibly prevent inhibitory development of the three subtypes of dyscalculia -preverbal, verbal and Arabic numbers. As this study has discovered, early numeracy developed in kindergarten is important in the development of arithmetic skills (Göbel et al., 2014).

In 2015 about 20% of the Norwegian 10th graders scored on level one in the PISA tests. According to OECD (2014), this indicates that these pupils are at risk of falling outside the community. In recent years there has been a focus on early intervention in the educational system, that can have yielded good results. In contrast to the 10th graders, the Norwegian 4th and 5th graders in 2015, scored the highest on TIMSS in mathematics of any Nordic country. Nevertheless, the kindergartens and schools can still improve by focusing even more on preventive work, since early abilities within mathematics are important, keeping in mind that numerical knowledge at age 7 can predict socioeconomic status at age 47 (Ritchie & Bates, 2013).

5.4 Conclusion and summary

The purpose of this study was to investigate to what extent number sense in 5-year-olds can predict arithmetic skills one year later. Based on previous studies, control variables were used - early numeracy, non-verbal intelligence and age. The results showed that ANS in kindergarten does not explain a significant amount of variance in arithmetic abilities one year later, taking the selected control variables into account. In other words, number sense at age 5 does not have a predictive role in arithmetic skills one year later. Instead, it was discovered that early numeracy in kindergarten is a unique predictor of arithmetic skills in 1st grade. In addition, it was hypothesized that ANS acuity at age 5 can predict ANS acuity one year later. This hypothesis was not met. In the analysis, it was found that number sense (K) from

Kindergarten had a weak and non-significant relation with number sense (G1) in 1st grade. This finding was surprising, and it can be a question whether the used dot comparison task does indeed tap the same skill at both time points. Furthermore, number sense (G1) in 1st grade had a very high positive relation with arithmetic, suggesting that both tests assessed almost the same skill, even though the tasks were very different. Another possible explanation could be that before children start school, ANS is a precursor skill that is influenced by, for example, the home environment. At school entry however, the children receive systematic instruction in arithmetic, and the relation between ANS and arithmetic becomes bidirectional: the arithmetic helps improve the ANS and the ANS helps improve the arithmetic. This could result in that the development of number sense is not linear, but becomes exponential after school entry, and therefore the relation from Kindergarten to 1st grade becomes weak.

A review of theory and empirical findings showed that the results in the study corresponds with previous findings to some extent. Several studies have found that early numeracy skills are important for the development of arithmetic skills, especially within the selected age group of this study (Bonny & Lourenco, 2013; Desoete et al., 2012; Göbel et al., 2014; Toll et al., 2015). So, the results of this study suggest that early numeracy is an important skill to develop towards the end of kindergarten. Required abilities within the three modules of the triple code model should receive focus from an early age (Dehaene, 1992) and aiming for the integration of all three codes is also essential. In summary, this concerns all skills from the ability to comprehend quantities, count, and knowledge of the Arabic numerals, to the relationships between the three representations of numbers, i.e., Arabic number '3' corresponds to the number word 'three', and represents the amount of three (***).

A closer examination of the results when compared with the empirical findings, suggests that to find ANS as a predictor for arithmetic, it needs to be measured at an earlier stage in development. Based on theory, an explanation for the findings could be that number words and Arabic numerals overlap with ANS, and thereby they become more important as predictors when the child has reached a certain level in its development. Nevertheless, ANS can be an important ability to test when there is a suspicion of difficulties in the acquisitions of early numeracy and arithmetic.

To conclude, ANS in Norwegian 5-year-old's is not a predictor of their arithmetic abilities one year later. It is the abilities within early numeracy that are crucial for Norwegian 5-year-olds' arithmetic abilities, one year later. Skills such as counting and knowledge of Arabic numerals should be strengthened to promote successful arithmetic development.

Difficulties in acquiring these skills should raise suspicion of possible future mathematical difficulties.

5.5 The way forward

Current research on number senses significance for mathematical development is divided. Some studies found that ANS has a relationship with mathematical ability and can be a predictor of early numeracy and arithmetic skills (Bonny & Lourenco, 2013; Libertus et al., 2011, Desoete et al., 2012; Libertus, et al., 2013; Mazzocco et al., 2011; Toll, et al., 2015). Other studies found that ANS's relation and predictive role, vanishes when controlling for other abilities, particularly those included within early numeracy (Bonny & Lourenco, 2013; Desoete et al., 2012; Göbel et al., 2014; Fuhs & McNeil 2013; Sasanguie et al., 2014; Toll et al., 2015). One recurring and crucial question from the empirical findings is whether it is age or the developmental stage that is decisive for ANS's predictive role. Further research is important for clarification of this issue.

Additionally, research on the mapping between symbolic and non-symbolic number sense would be interesting. In the early school years, pupils develop the ability to map symbolic representations onto non-symbolic quantities. This makes mapping an interesting intermediate skill that could form a bridge between early numeracy and arithmetic. Today, studies have found that the development of the mapping between symbolic and non-symbolic quantities are related to their math performance (Mundy & Gilmore, 2009; Toll et al., 2015). However, the causal relation is uncertain. Further studies on the mapping system would therefore be of interest in clarifying if an increased ability to manipulate symbolic representations results in an increased ability to map symbolic and non-symbolic representations, and if an increased mapping ability results in better symbolic number skills.

Finally, more research in number sense's role for the development of early numeracy is needed. Knowledge of this relation could tell us if a lack of skills within number sense can cause difficulties in acquiring early numeracy. If research shows that number sense does predict early numeracy, it would indicate that number sense indirectly predicts arithmetic skills, as early numeracy predicts arithmetic more directly.

References

Agrillo, C. (2015). Numerical and arithmetical abilites in non-primate species. In C. Kadosh, & Dowker, *The Oxford Handbook of numerical cognition* (pp. 214-236). Gostport, Hampshire: Oxford University Press.

American Psyhiatric Association (Ed,). (2013). *Diagnostic and stastical manual of mental disorders: DSM-5 (5th ed).* Washinton, D.C: American Psyhiatric Association.

Andersson, U., & Lyxell, B. (2007). Working memory deficits in children with mathematical difficulties: A general or specific deficit? *Journal of experminetal child psycholog, 96(3)*, pp. 197-228.

Arán, Filippetti, V., & Richaud , M. (2017). A structual equation modeling of executive functions, IQ and mathematical skills in primary students: Differential effects on number production, mental calculus and arithmetical problems. *Child Neuropsychology, 23(7)*, pp. 864-888.

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

Aunio, P., & Räsänen, P. (2015). Core numerical skills for learning mathematics in children aged five to eight years - a working model for educators. *European Journal of Early Childehood Education Research 24(5)*, pp. 684-704.

Aunola, K., Leskinen, E., Lerkkanen, M.-K., & Numri, J.-E. (2004). Developmental Dynamics of Math Performance From Preschool to Grade 2. *Journal of Educatinal Psychology, 96 (4)*, pp. 699-713.

Befring, E. (2015). *Forskningsmetoder i utdanningsvitenskap.* Oslo: Cappelen Damm akademisk .

Befring, E. (2015). *Forskningsmetoder i utdanningsvitenskap* . Oslo: Cappelen Damm akademisk.

Bergem, O. K., Kaarstein, H., & Nilsen, T. (2016). *Vi kan lykkes i realfag. Resultater og analyser fra TIMSS 2015.* Oslo: Universitet i Oslo IL.

Bonny, J., & Lourenco, S. (2013). The approximate number system and its relation to early math achievment: Evidence from preschool years. *Journal of Experimental Child Psychology, 114(3)*, pp. 375-388.

Brigestocke, S., Moll, K., & Wiebe, S. (2016). *TOBANS: Test of Basic Arithmetic and Numearcy Skills.*

Butterworth, B., Varma, S., & Laurillard, D. (2011). Dyscalculia: From Brain to Education. *Science, 322(6033)*, pp. 1049-1053.

Cook, T., Campbell, D., & Day, A. (1979). Quasi-experimentation: Design & Analysis issues for field settings. *(VOL. 351): Hoghton Mifflin Boston*.

De Forente Nasjoner. (1989). *Barnekonvensjonen.* http://barneombudet.no/wp-content/uploads/2013/09/fns_barnekonvensjon.pdf: Retrievd from .

De Vaus, D. (2014). *Surveys in social research (6th edition).* London: Routlegde.

Dehaene, S. (1992). Varieties of numerical abilites. *Cognition 44(1-2)*, pp. 1-42.

Dehaene, S. (2001). Précis of The Number Sense. *Mind & Language, 16(1)*, pp. 16-36.

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

Den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora (ED). (2016). *Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi. De nasjonale forkningsetiske komiteene.* https://www.etikkom.no/globalassets/documents/publikasjoner-som-pdf/60125_fek_retningslinjer_nesh_digital.pdf: Retreived from .

Deseoete, A., Ceulemans, A., De Weerdt, F., & Pieters, S. (2002). 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*, pp. 64-81. https://doi.org/10.1348/2044-8279.002002.

Desoete, A. (2015). Cognitive predictors of mathematical abilites and disabilities. In C. Kadosh, & Dowker, *The Oxford Handbook of numerical cognition* (pp. 915-930). Gotsport, Hampshire: Oxford University Press.

Desoete, A., Stock, P., Schepens, A., Baeyens, D., & Roeyers, H. (2009). Classification, Seriation, and Counting in Grades 1, 2, and 3 as Two-Year Longitdinal Predictors for Low Achieving in Numerical Facility and Arithmetical Achievemnt. *Journal of Psychoeducational Assessment, 27(3)*, pp. 252-264.

Desoete, A., Stock, P., Schepens, A., Baeyens, D., & Royers, H. (2009). Classificati, Seriation, and Counting in Grades 1, 2, and 3 as Tow-Year Longitudinal Predictors for Low Achieving in Numerical Facility and Arithmetical Achievement? *Journal of Psychoeducational Assessment, Volume 27, Number 3*, pp. 252-264.

DeWind, N., & Brannon, E. (2013, April). Malleability of the approximate number system, effects of feedback and training. *Frontiers in Human Neuroscience, Volume 6, Article 68*, pp. 1-10.

Ducan, G., Dowsett, C., Magnuson, K., Huston, A., Klebanov, P., & Japel, C. (2007). School readiness and later achievment. *Developmental Psychology, 43(6)*, pp. 1428-1446.

Dunn, L., Dunn, L., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale : Manusal (2. etg).* London: GL Assessment.

Field, A. (2009). *Discovering statistics using SPSS : (and sex and drugs and rock 'n' roll) (3. edition).* Los Angels: SAGE.

Fuhs, M., & McNeil, N. (2013). ANS acuity and mathematics ability in preschollers from low-income homes: contributions of inhibitory control. *Developmental Science, 16:1*, pp. 136-148.

Gall, M., Gall, J., & Borg, W. (2007). *Educational research: an introduciton (8th ed. utg.).* Boston, Mass: Allyn and Bacon.

Geary, D. (2000). From incancy to adulthood: the development of numerical abilites. *European Child & Adolscent Psychiatry, 9 (2)*, pp. II/11-II/16.

Geary, D. (2013). Early foundations for mathematics learning and their relations to learning disabilites. *Current directions in psychological science, 22(1)*, pp. 23-27.

Geary, D. (2015b). The Classification And Cognitive Characteristis of Mathematical Disabilities in Children. In R. Kadosh, & A. Dowker, *The Oxford Handbook of Numerical Cognition* (pp. 767-786). Gotsport, Hampshire: Oxford University Press.

Geary, D. C. (2015a). Preschoole children's quantiative knowldege and long-term risk for functional innumeracy. In S. J. Chinn, *The Routlede international handbook of dyscalculia and mathematical learning difficulteis* (pp. 232-242). Abindon, Oxon; New York, NY: Routlede/Taylor & Francis Group.

Geary, D., Hoard, M., Nugent, L., & Byrd-Craven, J. (2008). Development of Number Line Representations in Children with Mathematical Learning Disability . *Developmental Neuropsychology, 33(3)*, pp. 277-299.

Göbel, S., Watson, S., Lervåg, A., & Humle, C. (2014). Children's Arithmetic Development: It Is number Knowlede, Not the Approximate Number Sense, That Counts. *Psychological Sciense, 25(3)*, pp. 789-798.

Halberda, J., & Feigenson, L. (2008). Developmenal change in the acuity of "number sense". The approximate number system in 3-, 4.-, 5-, and 6-year-olds and adults. *Developmental Psychology, 44(5)*, pp. 1457-1465.

Hornigold, J. (2015). Learning Disabilites: mathematics characteristics and instructional exemplars. In S. J. (Ed), *The Routledge international handbook of dyscalculia and mathematical learning difficulties* (pp. 315-325). Abingdon, Oxon; New York; NY: Routledge/Taylor & Francis Group.

Hyde, D., Khanum, S., & Spelke, E. (2014). Breif non-symbolic, approximate number parctice enhances subsequent exact symbolic arithmetic in children. *Cognition, 131(1)*, pp. 92-107.

IMB Corp. (Releaed 2017). *IBM SPSS Statistics for Windows (Version 25.0).* Armonk, NY: IBM Corp.

Jordan, N., Kaplan, D., Nabors, O., & Locuniak, M. (2006). Number Sense Growth in Kindergarten: A Longitudinal Invetiagtion of Children at Risk for Mathematics Difficulties. *Child Development, 77(1)*, pp. 153-175.

Jordan, N., Kaplan, D., Ramineni , C., & Locuniak, M. (2009). Early Math Matters: Kindergarten Numer Competence and Later Mathematics Outcomes. *Developmental Psychology, 45(3)*, pp. 850-867.

Kleven, T. (2002a). Begrepsoperasjonalisering. In I. Lund, *Innføring i forskningsmetodologi* (pp. 141-183). Oslo: Unipub forlag.

Kleven, T. (2002b). Ikke-eksperimentell desing. In I. T. (Red), *Innføring i forskningsmetodologi (2. utg)* (pp. 265-285). Bergen: Fagbokforlaget.

Korhonen, J., Linnanmäki, K., & Aunio, P. (2014). Learning difficulties, academic well-being and educational dropout: Aperson-centerd approach. *Learning and Individual Differences, 31*, pp. 1-10.

Kunnskapsdepartementet. (2006). *... og ingen sto igjen. Tidlig innsats for livslang læring. St.Meld. 16 (2006-2007).* pp. 11-12. Oslo. Retrieved from https://www.regjeringen.no/contentassets/a48dfbadb0bb492a8fb91de475b44c41/no/pdfs/stm200620070016000dddpdfs.pdf: Kunnskapsdepartementet.

Kunnskapsdepartementet. (2017). *Rammeplan for barnehagen: Antall, rom og form.* pp. 53-54. Oslo: Retrievd from https://www.udir.no/globalassets/filer/barnehage/rammeplan/rammeplan-for-barnehagen-bokmal2017.pdf: Kunnskapsdepartementet.

Libertus, M., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Sciense 14:6*, pp. 1292-1300.

Libertus, M., Feigenson, L., & Halberda, J. (2013). Is approximate number precision a stable predictor of math ability? . *Learning and Individual Differences, 25*, pp. 126-133.

Lund, T. (2002). Metodiske prinsipper og referanserammer. In K. T.-A. I. K. T. A Lund Thorleif, *Innføring i forskningsmedodologi (2. utg)* (pp. 79-123). Bergen: Fagbokforlaget.

Lyster, S.-A., Horn, E., & Rygvold, A. (2010). Ordforråd og ordforrådsutvikling hos norske barn og unge. Resultater fra en utprøving av Britich Picture Vocabulary Scale Second Edition (BPVS II). *Spesialpedagogikk*, pp. 35-43.

Mazzocco, M., Feigenson, L., & Halberda, J. (2011). Preschollers' Precision of the Approximate Number System Predicts Later School Mathematics Performance. *PLoS ONE 6(9)*, pp. 1-8.

McCring, K., & Bridsall, W. (2015). Numerical abiliteis and arithmetic in infancy. In C. Kadosh, & Dowker, *The Oxford Handbook of numerical cognition* (pp. 258-274). Gosport, Hampshire: Oxford University Press.

Mundy, E., & Gilmore, C. (2009). Children's mapping between symbolic and nonsymbolic representations of number. *Journal of Experimental Child Psychology, 103(4)*, pp. 490-502.

Neider, A., & Dehaene, S. (2009, April 20). Representation of Number in the Brain. *Annual Review of Neurosciense , 32*, pp. 185-208.

Obersteiner, A., Reiss, K., & Ufer, S. (2013). How training on exact or approximate mental representations of numer can enhance first-grade students' basic number processing and arithmetic skills. *Learning and Instruction, 23*, pp. 125-135.

OECD. (2014). *PISA 2012 results: what studtens know and can do: Student performance in mathematics, reading and science.* Paris: OECD Publiching.

OECD. (2016). *PISA 2015 Results (Volume I): Exellence and Equity in Education.* Paris: OECD Publiching.

Park, J., & Brannon, E. (2013). Training the Approximate Number System Improves Math Proficiency. *Psycological Science, 24(10)*, pp. 2013-2019.

Passolunghi, M., & Lanfranchi, S. (2012). Domain-specific and domain-general precursors of mathematical achievement: A longitudial study from kindergarten to first grade. *British Journal of Educational Psychology, 82(1)*, pp. 42-63.

Piazza, M., Pica, P., Izard, V., Spelke, E., & Dehaene, S. (2013). Education Enhances the Acuity of the Nonverbal Approximate Number System. *Psychological Science, 24(6)*, pp. 1037-1043.

Price, D., & Ansari , D. (2013). Dyscalculia; Characteristics, Causes and Treatments. *Numeracy 6 (1) http://dx.doi.org/10.5038/1936-4660.6.1.2*.

Raven, J. C. (1998). *The Raven's Progressive Matrices.* Oxford Psychologist Press Oxford.

Ritchie, S., & Bates, T. (2013). Enduring Links From Childhoos Mathematics and Reading Achievment to Adult Socioeconomic Status. *Psychological Sciensce, XX(X)* , pp. 1-8.

Sasanguie, D., Göbel, S., Moll, K., Smets , K., & Reynvoet, B. (2013). Approximate number sense symbolic number processing, or number-space mappings: What underlies mathematics achievemnt? *Journal of Experimental Child Psychology, 114(3)*, pp. 418-431.

Schmithorst, V., & Brown, R. (2004). Empirical validation of the triple-code model of numerical processing for complex math operations using functional MRI and group Independent Component Analysis of the mental addition and subtraction of fractions. *NeuroImage, 22(3)*, pp. 1414-1420.

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

Siegler, R., & Braithwaite, D. (2004). Numerical Development. *Annual Review of Neuroimage, 22(3),* pp. 1414-1420.

Starkey, P. (1992|). The eraly development of numerical reasoning. *Cognition, 43(2)*, pp. 93-126.

Stock, P., Desoete, A., & Royers, H. (2009). Detecting Childrn With Arithmetic Disabilites From Kindergarten: Evidence From a 3-Year Longitudinal Study on the Role of Prepartory Arithmetic Abilites. *Journal of Learning Disabilites, 43(3)*, pp. 250-268.

Storvik, L. (2019). Barn fra barnehager med god kvalitet får et års forsprang på skolen . *Utdanningsnytt*, pp. Retrievd from https://www.utdanningsnytt.no/forste-steg/artikler/2019/mai/barn-fra-barnehager-med-god-kvalitet-far-et-ars-forsprang-pa-skolen/.

Toll, S., & Van Luit, J. (2013). The development of early numeracy ability in kindergartners with limited working memory skills. *Learning and Individual Differences, 25*, pp. 45-54.

Toll, S., Van Viersen, S., Kroesbergen, E., & 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*, pp. 10-17.

Udaninngsdirektoratet. (2017, December 21). Rammeverk for grunnlegende ferdigheter. *Utdanningsdirektoratet*, p. 2. Retrieved from https://www.regjeringen.no/contentassets/a48dfbadb0bb492a8fb91de475b44c41/no/pdfs/stm200620070016000dddpdfs.pdf.

Von Aster, M. (2000). Developmental cognitive neuropsychology of number processing and calculation: varieties of developmental dyscalculia. *European Child & Adolescent Psychairty, Volume 2, supplement 2*, pp. 41-57.

Von Aster, M., & Shalev, R. (2007). Numer development and developmental dyscalculia. *Developmental Medicine & Child Neurology, 49(11)*, pp. 868-873.

Wang, J., Odic, D., Halberda, J., & Feigenson, L. (2016). Changing the precision of preschoolers' approximate number system representations changes their symbolic math performance. *Journal of Experimental Child Psychology, 147*, pp. 82-99.

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

Appendix A

Histograms of the distribution of the original variables.

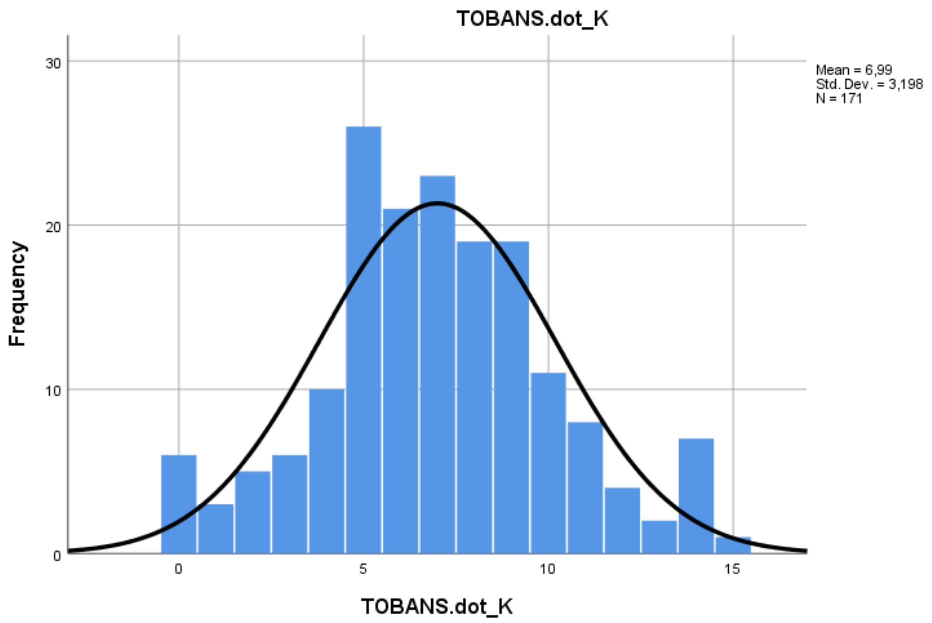

Figure 2 *Histogram of the Distribution to Dot Comparison (K)*

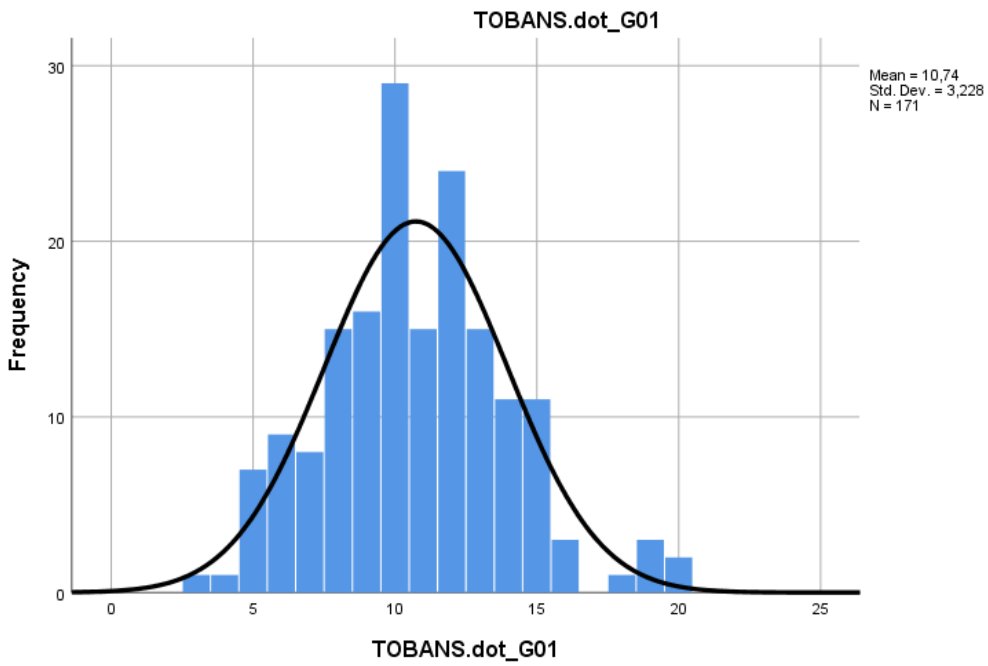

Figure 3 *Histogram of the Distribution to Dot Comparison (1st)*

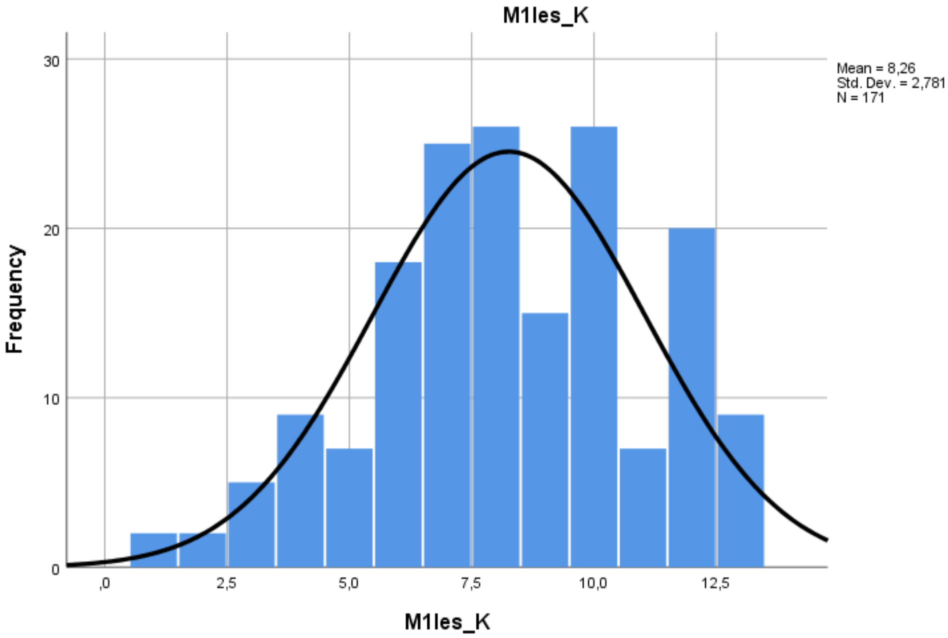

Figure 4 *Histogram of the Distribution to number reading*

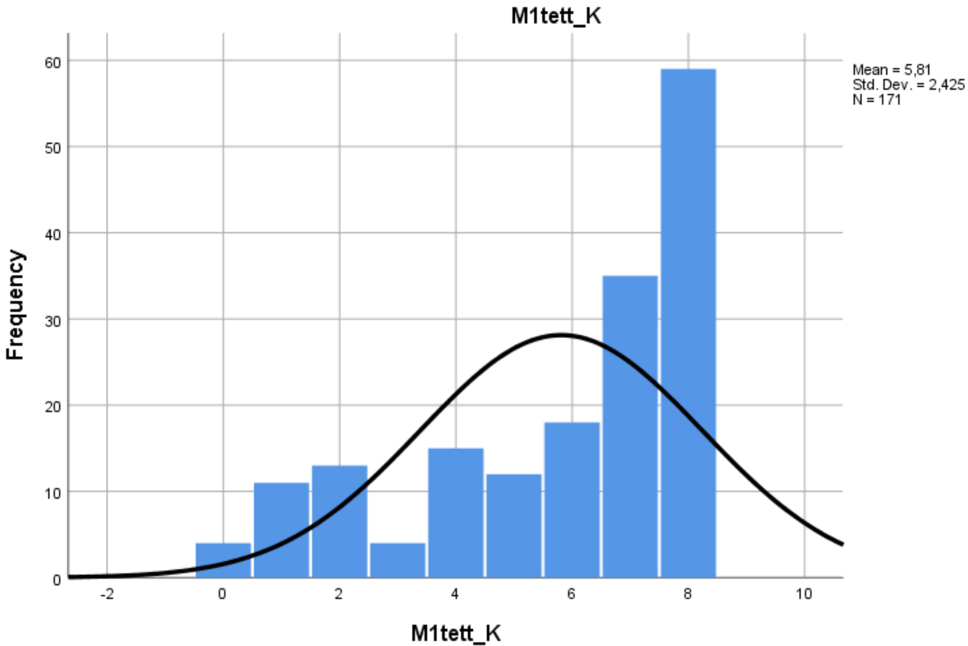

Figure 5 *Histogram of the Distribution to counting on*

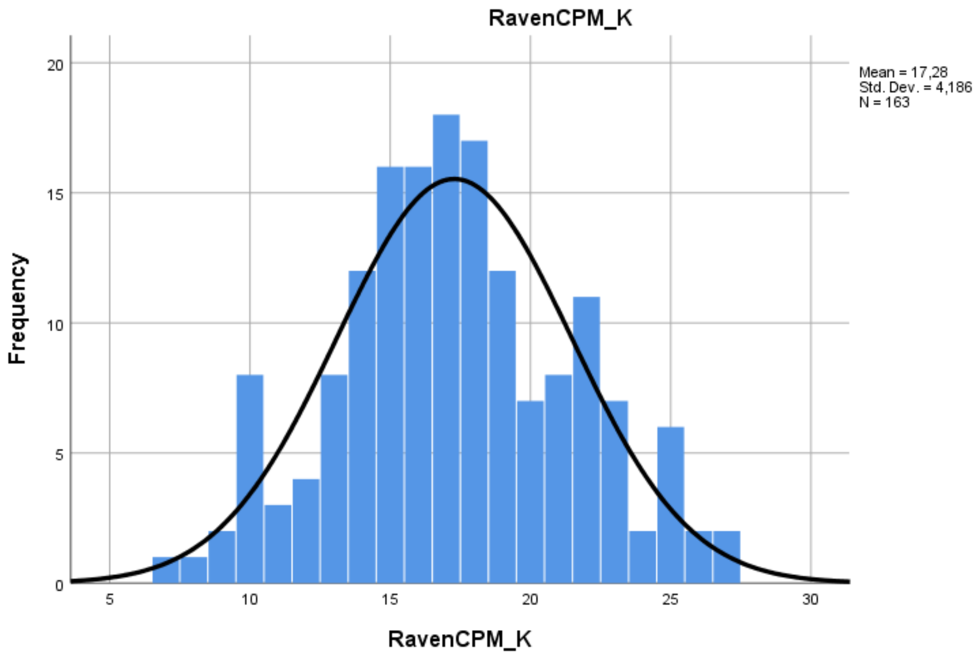

Figure 6 *Histogram of the Distribution to Raven*

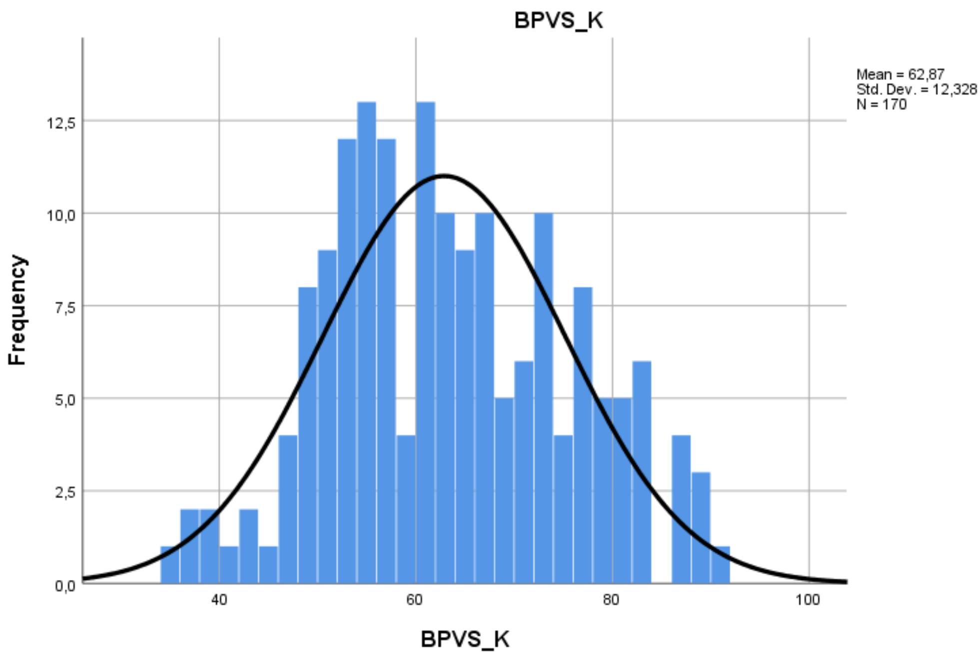

Figure 7 *Histogram of the Distribution to BPVS*

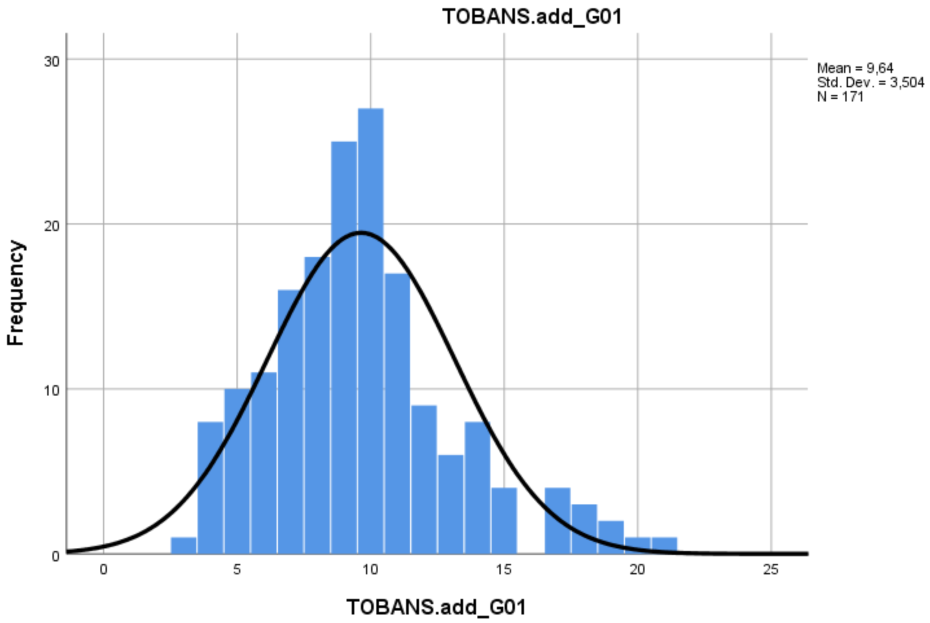

Figure 8 *Histogram of the Distribution to addition*

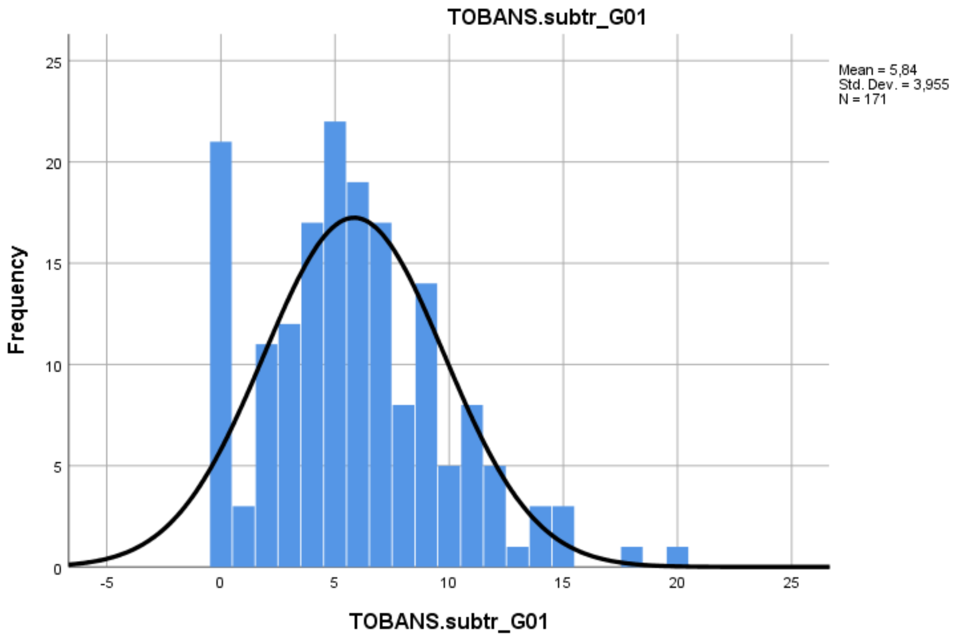

Figure 9 *Histogram of the Distribution to subtraction*

Appendix B

Historams of the combined variables

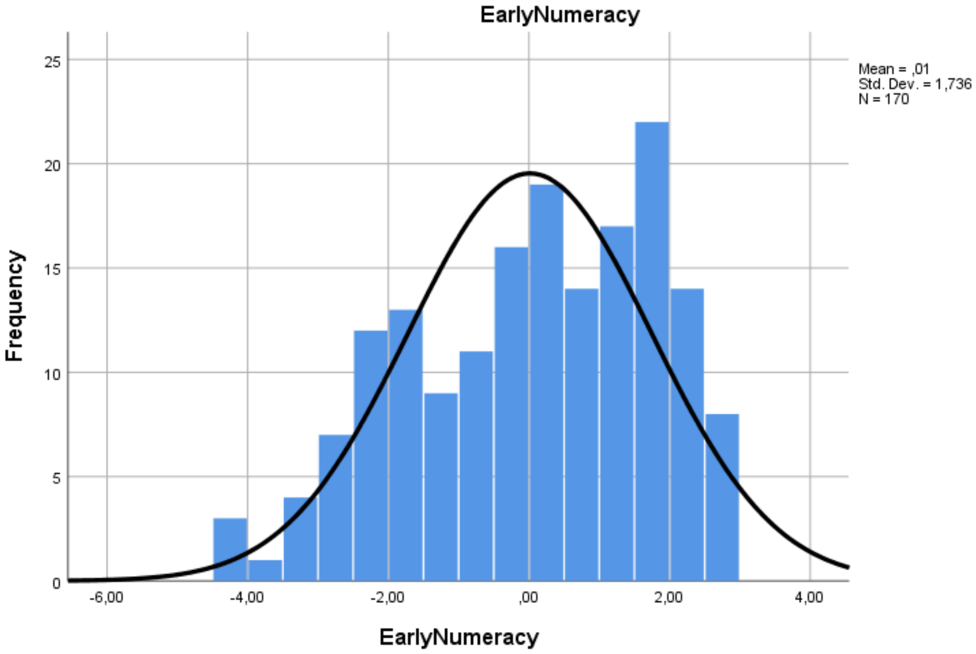

Figure 10 *Histogram of the Distribution to Early Numeracy*

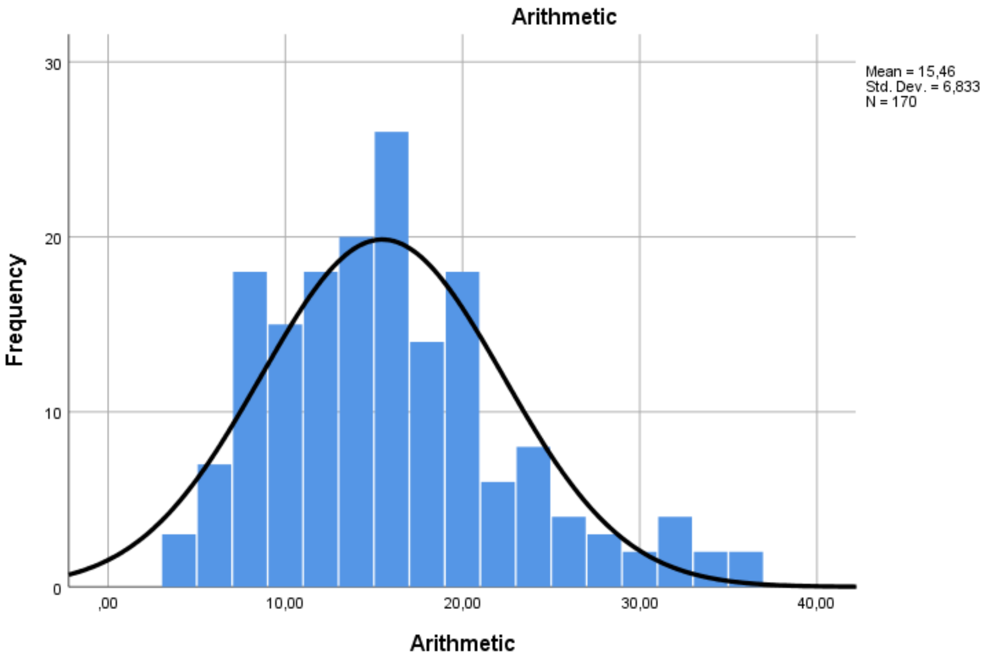

Figure 11 *Histogram of the Distribution to Arithmetic*

Appendix C

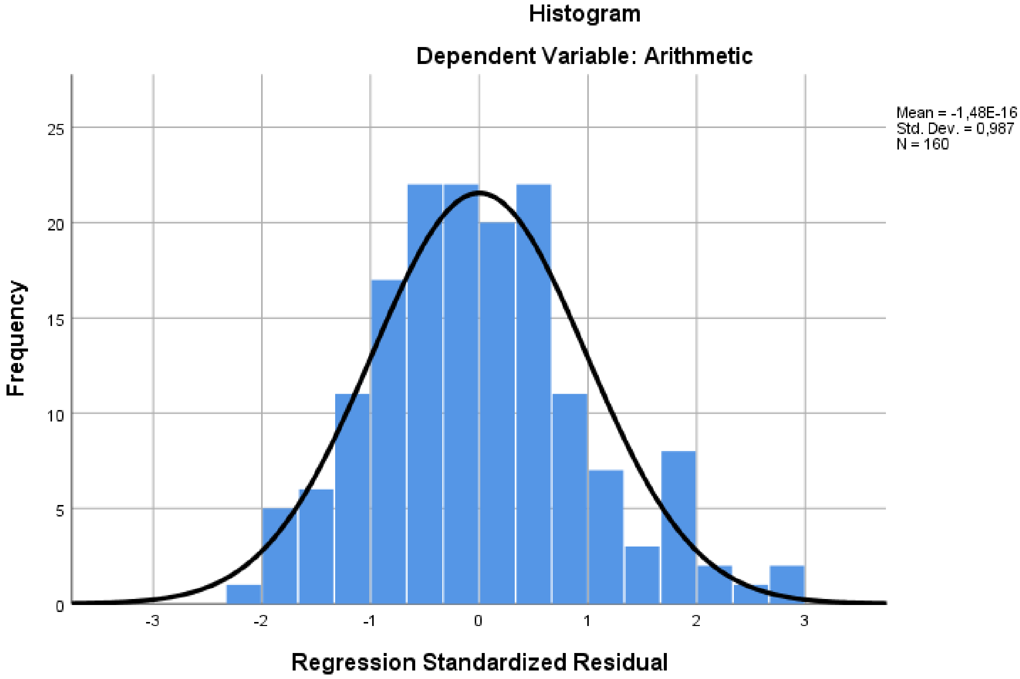

Figure 12 *Histogram for the distribution of expected residual values. Arithmetic*

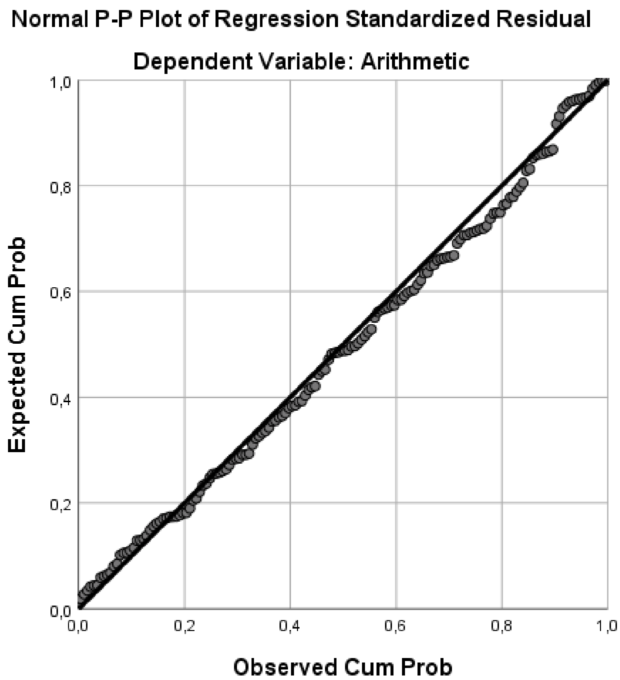

Figure 13 *P-P plots of regression standarized residual*

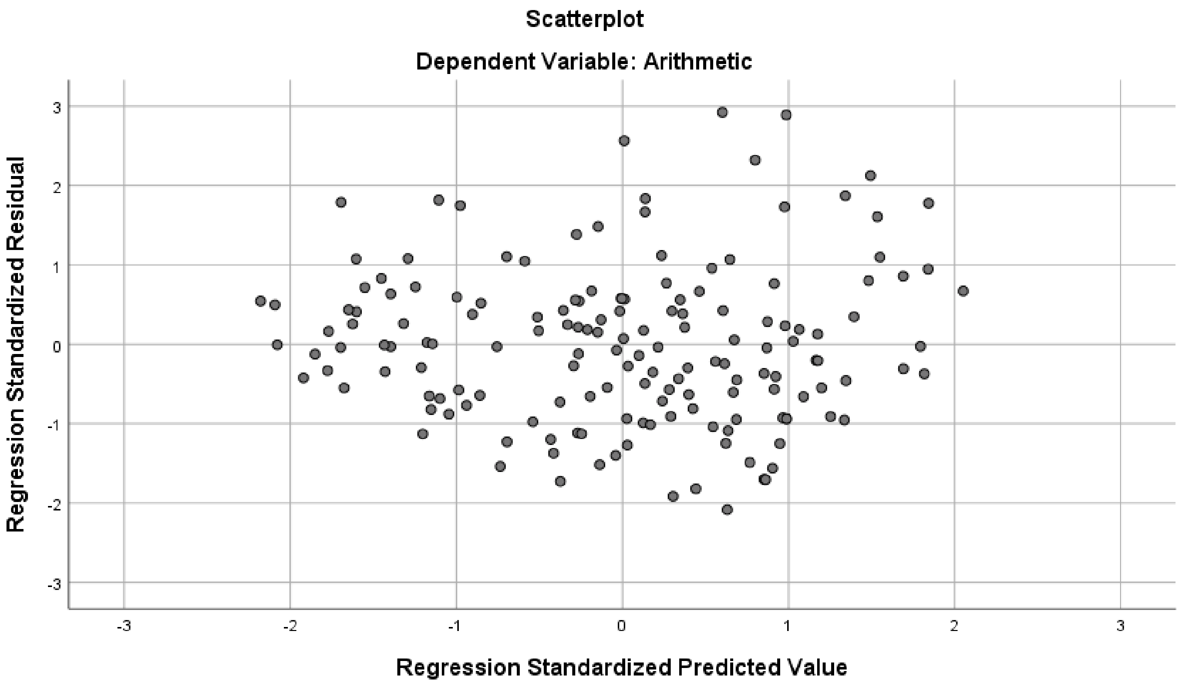

Figure 14 *Scatterplot of the distribution of residuals of expected vaules: Arithmetic*

Partial Regression Plot

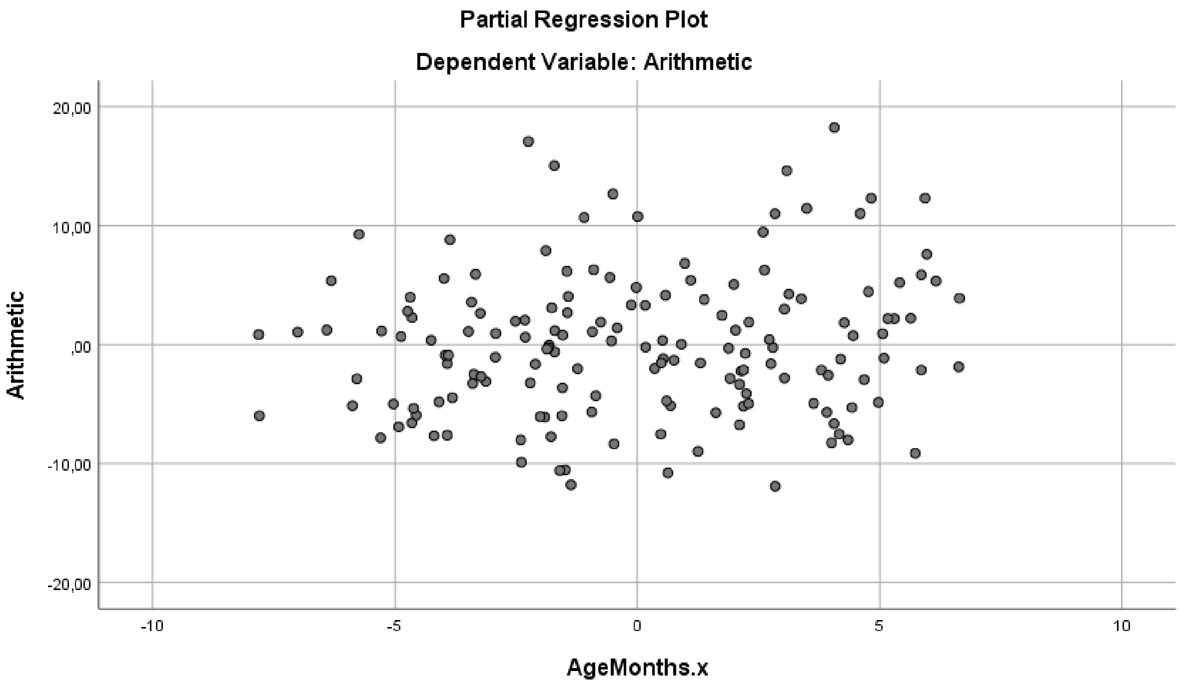

Figure 15 *Partial Regression Plot Arithmetic - age*

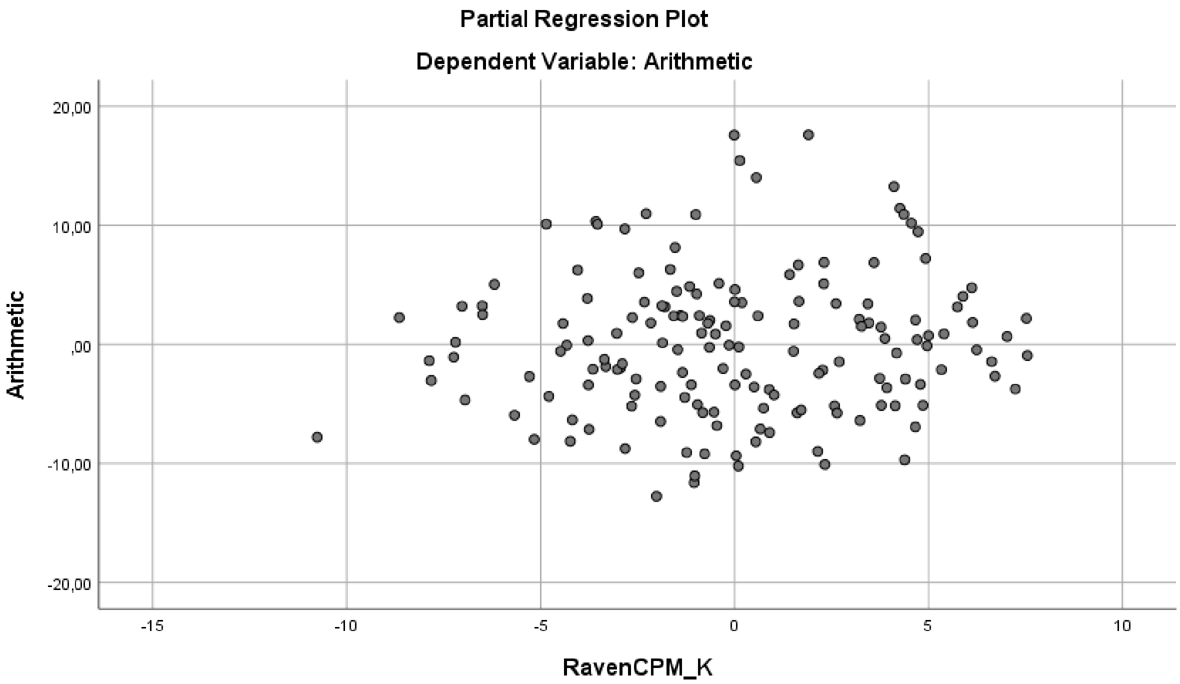

Figure 16 *Partial Regression Plot Arithmetic - raven*

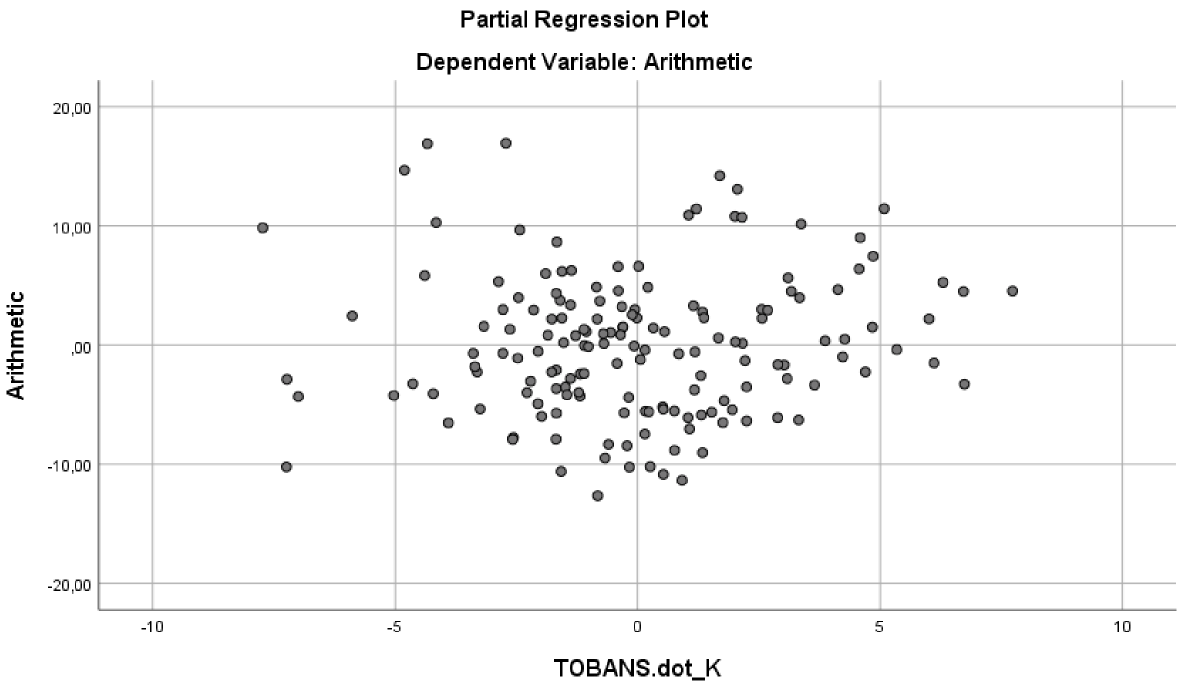

Figure 17 *Partial Regression Plot Arithmetic - Dot comparison (K)*

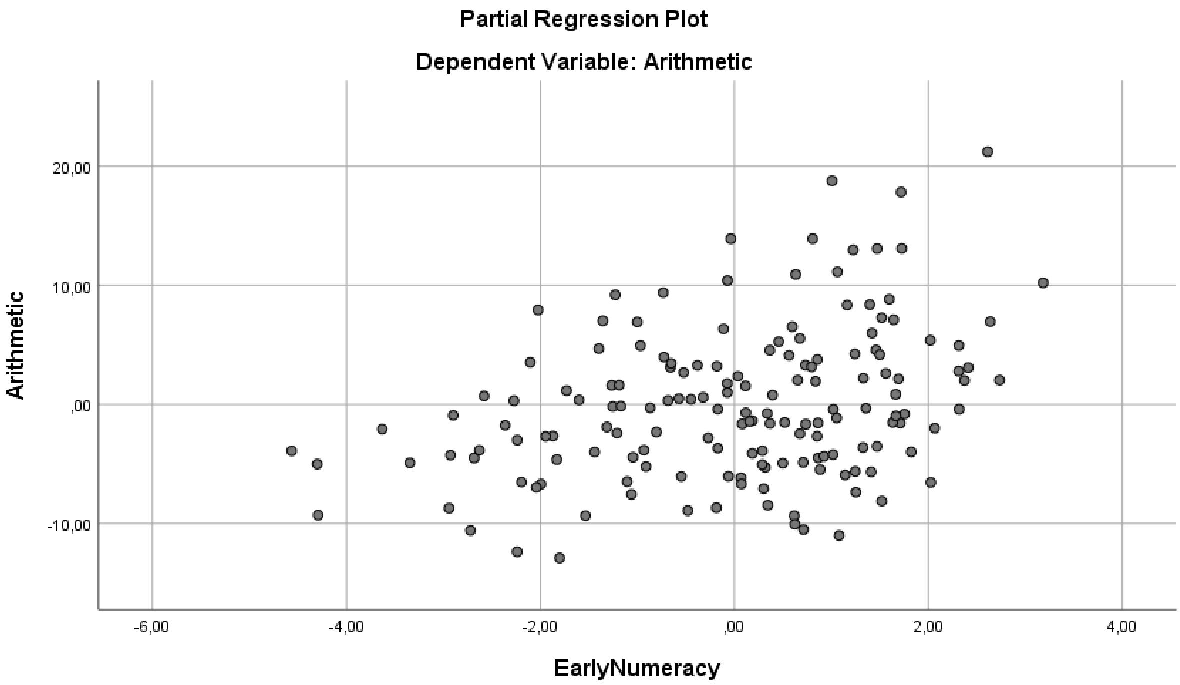

Figure 18 *Partial Regression Plot Arithmetic - early numeracy*