Journal of Experimental Child Psychology 219 (2022) 105390

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

Mathematical skills of 11-year-old children born very preterm and full-term

Hilma Halme [a,*], Jake McMullen [a], Cristina E. Nanu [a], Anna Nyman [b], PIPARI Study Group [c,1], Minna M. Hannula-Sormunen [a]

[a] Department of Teacher Education, 20014 University of Turku, Finland
[b] Department of Psychology, 20014 University of Turku, Finland
[c] Turku University Hospital, 20521 Turku, Finland

ARTICLE INFO

Article history:
Received 15 December 2020
Revised 18 November 2021
Available online 23 February 2022

Keywords:
Mathematical skills
Primary school
Preterm
Spontaneous focusing on numerosity
Spontaneous focusing on quantitative relations
Rational number

ABSTRACT

Preterm birth affects the academic development of children, especially in mathematics. Remarkably, only a few studies have measured specific effects of preterm birth on mathematical skills in primary school. The aim of this study was to compare 11-year-old children, with an IQ above 70, born very preterm (N = 64) and full-term (N = 72) on a variety of 5th grade mathematical skills and cognitive abilities important for mathematical learning. The measures were spontaneous focusing on numerosity (SFON), spontaneous focusing on quantitative relations (SFOR), arithmetic fluency, mathematics achievement, number line estimation, rational number magnitude knowledge, mathematics motivation, reading skills, visuospatial processing, executive functions, and naming speed. The children born very preterm and full-term differed in arithmetic fluency, SFON and SFOR. Domain general cognitive abilities did not fully explain the group differences in SFON and SFOR. Retrospective comparisons of the samples at the age of five years showed large group differences in early mathematical skills and cognitive abilities. Despite lower early mathematical skills, the children born very preterm reached peer equivalent performance in many mathematical skills by the age of 11 years. Nevertheless, they appear less likely to focus on implicit mathematical features in their everyday life.

* Corresponding author.
E-mail addresses: hilma.halme@utu.fi (H. Halme), jamcmu@utu.fi (J. McMullen), eucrna@utu.fi (C.E. Nanu), minna.hannula-sormunen@utu.fi (M.M. Hannula-Sormunen).

[1] A full list of the members of the Pipari Study Group is in the Acknowledgments.

https://doi.org/10.1016/j.jecp.2022.105390
0022-0965/

Introduction

Over the last few decades, preterm birth survival rates have been increasing and a vast number of studies report adverse neurodevelopmental effects associated with preterm birth, especially amongst children born very preterm or with a very low birth weight (Aarnoudse-Moens, Weisglas-Kuperus, Van Goudoever, & Oosterlaan, 2009; Sansavini, Guarini, & Caselli, 2011; Taylor, Espy, & Anderson, 2009). There is a high prevalence of learning difficulties amongst children born preterm compared to full-term that are most pronounced in mathematics (Sansavini et al., 2011; Taylor et al., 2009; Twilhaar, de Kieviet, Aarnoudse-Moens, van Elburg, & Oosterlaan, 2018). While meta-analyses have shown that arithmetic scores and mathematics achievement are lower in children born preterm than full-term (Aarnoudse-Moens et al., 2009; Twilhaar et al., 2018), mathematical skills of children born preterm have been commonly measured using a single composite achievement test. This results in limited knowledge on what specific mathematical skills are affected by preterm birth (Johnson et al., 2009). In particular, there is little evidence regarding how preterm birth affects more advanced mathematical skills, which are important for applying mathematical knowledge in future learning, everyday life, and future careers (ACME, 2011). Studies have suggested that the lower mathematical performance of children born preterm is due to weaker cognitive abilities (Simms et al., 2013, Taylor et al., 2009). Thus, the present study aims to examine if 11-year-old children born very preterm have lower mathematical performance than children born full-term on a wide range of specific mathematical skills, including more advanced mathematical skills. Furthermore, it will investigate if cognitive abilities explain group differences in these different mathematical skills.

Mathematical skills in children born preterm

Meta-analyses show evidence of disadvantages faced by children born preterm in learning mathematical skills (Aarnoudse-Moens et al., 2009; Twilhaar et al., 2018). For instance, Twilhaar et al. (2018) showed that standardized academic scores of children born preterm at five years of age or older were 0.71SD below full-term peers on arithmetic. Furthermore, birth weight and gestational age are positively correlated with mathematical abilities in children born preterm (Aarnoudse-Moens et al., 2009; Taylor et al., 2009; Wolke et al., 2015). In a very large cohort study of almost 1500 children born preterm, Wolke et al. (2015) showed that mathematical performance measured at eight years of age was significantly related to gestational weeks.

However, a review of the literature reveals inconsistencies in findings of lower mathematical performance both within and between specific mathematical skills, especially at school age. For example, weaker arithmetic fluency has been found in 14-year-old children born preterm with an extremely low birth weight (<1000 g; Litt et al., 2012). This suggests that the lower arithmetic fluency found in children born preterm in earlier school years (Clark & Woodward, 2015; Pritchard et al., 2009; Short et al., 2003) may persist over time. Consistent with weaker arithmetic fluency, studies have found weaker counting sequence knowledge and digit knowledge in primary (Korpipää et al., 2019; Pritchard et al., 2009; Simms et al., 2015) and secondary school (Clayton et al., 2021) among children born preterm. In contrast, Rose, Feldman, and Jankowski (2011) found that 11-year-old children born preterm do not have lower arithmetic fluency, but they have lower performance on applied problems. As there are still very few studies measuring specific mathematical skills in children born preterm, persistence of lower performance in different specific skills is still unclear.

Studies on specific mathematical skills seem to only be consistent with the finding that children born very preterm reach peer equivalent accuracy in estimating the magnitude of numbers by primary school (Clayton et al., 2021; Guarini et al., 2014; Guarini, Tobia, Bonifacci, Faldella, & Sansavini, 2020; Simms et al., 2015). For instance, there were differences in accuracy on symbolic and non-symbolic whole number comparison at the age of six years, but no longer at the age of eight years (Guarini et al., 2014; Guarini, Tobia, Bonifacci, Faldella, & Sansavini, 2020). Likewise, whole number line estimation at the age of 11 years may only be weaker in children born extremely preterm (<26 weeks; Simms et al., 2013) and not in very preterm (<32 weeks) after the age of eight years (Clayton et al., 2021; Simms et al., 2015). Even though children born preterm may have sufficient whole number magnitude processing by primary school, more research is needed to assess whether challenges may occur in advanced mathematical skills, such as rational number magnitude knowledge.

Children born preterm may have disadvantages in learning more complex mathematical skills, even if they would be capable of achieving peer equivalent performance on more basic mathematical skills. Weaker cognitive abilities may cause problems in learning complex mathematical content due to the increase in cognitive requirements with increasing content complexity (Clayton et al., 2021; Simms et al., 2015; Wehrle et al., 2016). Accordingly, rational number magnitude knowledge may prove more challenging than whole number magnitude knowledge for children born preterm. As rational numbers are a stepping-stone towards algebra and later mathematics, poor rational number knowledge could create a bottleneck in learning (DeWolf, Bassok, & Holyoak, 2015; Siegler et al., 2012). For instance, lower performance on algebra was found in children born very preterm (Clayton et al., 2021). Yet the effect of preterm birth on rational number knowledge has not been specifically examined.

Due to differences in cognitive demands, domain general cognitive abilities may also have different impacts on the development of different mathematical skills. There is evidence of mathematical skills of children born preterm being related to cognitive abilities, such as processing speed, executive functions, and visuospatial skills (Litt et al., 2012; Mulder, Pitchford, & Marlow, 2010; Rose et al., 2011; Sansavini et al., 2011; Simms et al., 2015, 2013; Taylor et al., 2009; Twilhaar, De Kieviet, Van Elburg, & Oosterlaan, 2020). Simms et al. (2015) found that basic numerical skills of children born very preterm aged 8–10 years were associated with lower working memory and visuospatial processing. Guarini et al. (2014, 2020) showed that children born very preterm without differences in full-scale IQ had persistently slower reaction times, when comparing non-symbolic magnitudes at the ages of eight and ten years. The slower reaction times were attributed to slower processing speed (Guarini et al., 2020). Litt et al. (2012) showed that sustained attention, visual memory and visuospatial working memory are related to mathematical calculation fluency in 14-year-old children born preterm. Thus, the current study also examines whether domain general cognitive abilities explain the potential differences in mathematics domain specific skills.

Spontaneous mathematical focusing tendencies

Mathematical skills are not needed only in mathematics lessons at school, but also in everyday situations. The frequency with which children spontaneously use their existing mathematical skills in activities that are not explicitly mathematical predicts their mathematical development (for recent review see McMullen, Chan, Mazzocco, & Hannula-Sormunen, 2019; Verschaffel et al., 2018). Most prominently, Spontaneous Focusing On Numerosity (SFON) is defined as the spontaneous (i.e. unguided) focusing of attention on exact numerosity and the use of exact numerosity in situations that are not explicitly mathematical (Hannula and Lehtinen, 2005). For example, in a SFON task involving making similar pictures, some preschool children notice by themselves the exact number of "spikes" the tester stamped on the back of a dinosaur and stamp exactly the same number of spikes. The others focus on other, non-numerical, aspects of the task, such as the location of the spikes, and stamp a random number of spikes on the dinosaur (Hannula and Lehtinen, 2005). SFON tendency, measured across several SFON tasks, is proposed to indicate a child's self-initiated practice of enumeration skills in everyday life (Hannula, Lepola & Lehtinen, 2010) and it supports the development of numeracy skills.

Individual differences in preschool SFON are related to both concurrent and later mathematical skills. For example, SFON tendency is a domain specific-predictor of arithmetic fluency and whole number line estimation (Hannula et al., 2010; Nanu et al., 2018), which may be weaker in children born preterm (Pritchard et al., 2009; Simms et al., 2013). However, there are very few published studies on SFON tendency measured in school-aged children. Kucian et al. (2012) found that primary school children between the ages of 7 and 11 years with diagnosed mathematical difficulties have lower SFON tendency than their cognitively matched peers. Hence, lower mathematical skills in 11-year-old children born very preterm could be similarly associated with lower SFON tendency.

The findings of the importance of SFON tendency in early mathematical development has led to research on a similar tendency of spontaneous focusing on quantitative relations (SFOR). SFOR is defined as the spontaneous recognition and use of exact quantitative relations in a mathematically unspecified situation (McMullen, Hannula-Sormunen, & Lehtinen, 2014). For example, noticing that there are twice as many oranges as apples in a bowl. Like SFON tendency, a higher SFOR tendency is proposed to facilitate mathematical learning through self-initiated practice with quantitative relations in everyday situations. SFOR tendency is an important developmental contributor of rational number knowledge (McMullen, Hannula-Sormunen, Laakkonen, & Lehtinen, 2016; Van Hoof et al., 2016). No previous research has examined SFOR tendency in children born preterm.

Previous mathematics related studies in the PIPARI research project

The present study is part of the multidisciplinary, longitudinal research project PIPARI (Development and functioning of Very Low Birth Weight Infants from Infancy to School Age), which follows children born preterm and full-term from birth to adulthood (https://sites.utu.fi/pipari/en). In a previous study of the current subsample of 11-year-old children born preterm and full-term, Hannula-Sormunen et al. (2017) examined the children's mathematics-related knowledge and skills at the age of five years (i.e. before primary school). At the age of five years, the mathematical sub-skills measured included digit knowledge, SFON, arithmetic, counting and geometry skills. In a latent profile analyses of mathematical skills, more profiles were needed for reliable grouping of the children born preterm compared to full-term due to higher variability in their mathematical sub-skills. Furthermore, more children born preterm belonged to below average skill profiles. Differences in the mathematical skill profiles of the children born preterm were associated with naming speed, phonological processing, and visuospatial working memory. Notably, this study did not explicitly compare the groups' overall mean values for different mathematical and cognitive skills at the age of five years. In addition, with a partly overlapping sub-sample, two studies showed lower digit knowledge, counting sequence knowledge, basic arithmetic skills (i.e. addition and subtraction), phonological processing and naming speed in children born preterm compared to full-term at the beginning of Grade 1 (Alanko et al., 2017; Korpipää et al., 2019). Thus, within this project, lower basic mathematical skills have been found before and at the start of school in children born preterm.

Previous studies in this project have also examined children's cognitive abilities at the age of 11 years. Nyman et al. (2017) found that children born very preterm, even without neurosensory impairment, had lower performance on the four domains of the WISC-IV compared to normative tests scores. Likewise, children born with a very low birth weight ($\leq$1500 g), even with normal cognitive development, showed lower performance on measures of working memory and visuospatial short-term memory, but not verbal short-term memory, compared to normative test means (Korpela et al., 2018). However, teacher and parent rated executive functions suggest fewer problems in 11-year-old children within this cohort than has been reported in other cohorts (Nyman et al., 2019b). If there are weaknesses in cognitive abilities, especially in working memory and visuospatial skills, these could be a cause for mathematical difficulties. To examine this, we will measure whether lower mathematical skills are due to weaker cognitive abilities when compared to a control group.

Despite the wealth of knowledge about the differences in children born preterm and full-term in the PIPARI research project, the only study examining group differences in mathematical performance in late primary school at the age of 11 years have used teacher-ratings and not examined specific mathematical skills. Nyman et al. (2019a) showed that 11-year-old children born very preterm with an IQ above 70 did not differ from their full-term born peers in teacher-rated mathematical performance. The lack of significant differences in teacher-ratings is surprising and encouraged more detailed examination of more objective measures of mathematical skills and knowledge.

Current study

The current study examines the performance of 11-year-old children born very preterm and full-term on a wide range of mathematical skills and mathematics-related domain general cognitive abilities. The primary aim is to determine whether lower mathematical performance is present at the age of 11 years in children born very preterm without severe cognitive deficits (i.e. IQ < 70). Importantly, we extend the research on specific mathematical skills in children born very preterm from basic skills and general math achievement to include also more advanced mathematical skills and spontaneous mathematical focusing tendencies. Thus we measure a comprehensive set of mathematical skills covering developmentally relevant aspects of number concept, including (a) basic skills of arithmetic fluency and whole number line estimation (Litt et al., 2012; Rose et al., 2011; Simms et al., 2015, 2013), (b) more advanced formal mathematical knowledge, in the form of curriculum-based mathematics achievement and rational number magnitude knowledge (Twilhaar et al., 2018; Van Hoof, Janssen, Verschaffel, & Van Dooren, 2015), and (c) SFON and SFOR, which are related to these basic and advanced mathematical skills (McMullen et al., 2019). The chosen mathematical skills also match the 5th grade (i.e. late primary school) curriculum and the typical mathematical development at the age of 11 years. In order to determine whether group differences are specific to the mathematics domain, we also included a basic reading skill and mathematics motivation.

The secondary aim is to examine whether lower domain general cognitive abilities account for group differences in specific mathematical skills between children born very preterm and full-term. Based on previous studies, the domain general cognitive abilities expected to be related to mathematical skills and knowledge in children born very preterm are naming speed (Hannula-Sormunen et al., 2017; Rose et al., 2011), visuospatial processing (Hannula-Sormunen et al., 2017; Clayton et al., 2021; Litt et al., 2012; Simms et al., 2015), and executive functions (Clayton et al., 2021; Mulder et al., 2010; Rose et al., 2011). Gender and maternal education are also considered, as both demographic variables may affect mathematical attainment in children born preterm (Doyle et al., 2015; Johnson & Breslau, 2000; Nyman et al., 2017; Wolke et al., 2015). Last, we conduct an ad-hoc retrospective analysis on the mathematical skills and cognitive abilities of these participants at the age of five years to examine what group differences existed already then.

Methods

Participants

The current PIPARI research project sample includes Finnish speaking 11-year-old children born very preterm (n = 67) and full-term (n = 72), who were in the 5th grade. The children had not been retained at any grade. All children in the current sample had participated in mathematics-related testing sessions at the age of five years (+0–2 months) before primary school. In Finland, children start primary school in the fall of the year when they turn seven years. Primary school has six grades. As the current study examines children with a full-scale intelligence quotient (FSIQ) within the average range (>70), three children born very preterm were excluded from the current analysis due to a FSIQ score less than 70 (n = 1) or lack of FSIQ data (n = 2) at the age of 11 years. The mean (SD) FSIQ of the 11-year-old children born very preterm was 94.12 (14.0). FSIQ data is not available for the children born full-term at the age of 11 years, as only the children born preterm are systematically followed up. Notably, the children born very preterm in the current study did not significantly differ in FSIQ from the PIPARI research project sample of 11-year-old children born preterm, who were in the 5th grade, but did not take part in this study (n = 76). Thus, this sample is representative of 11-year-old children born preterm who are in the 5th grade.

For all children, the exclusion criteria at five years of age were congenital anomalies, an IQ score below 70, or a diagnosed developmental disorder. The children born very preterm were born at Turku University Hospital in 2002–2006 with either a gestational age less than 32 weeks (n = 55) or a birth weight less than 1501 g (n = 60). The children born full-term were recruited by asking for participation from the first healthy boy and first healthy girl born each week at Turku University Hospital in 2002–2004. The inclusion criteria for the children born full-term were a gestational age of at least 37 weeks at birth, a birth weight higher than −2 SD of the Finnish growth chart norms, and Finnish as the mother tongue. The exclusion criteria for the children born full-term were congenital anomalies or syndromes, admission to neonatal care during the first week of life or a self-reported

Table 1
Demographic information and neonatal characteristics of the children born very preterm and full-term.

	Very preterm (n = 64)	Full-term (n = 72)	*p*
Age (years) mean (SD) [min, max]	11.81 (0.28) [11.30, 12.36]	11.76 (0.29) [11.29, 12.33]	$t(134) = 0.89$, $p = 0.38$
Gestational age (weeks) mean (SD) [min, max]	29.27 (2.57) [24, 35.29]	40.07 (1.20) [37.29, 42.14]	
Birth weight (g) mean (SD) [min, max]	1138 (305) [565, 1670]	3623 (441) [2850, 4630]	
Female *n* (%)	26 (40.6)	40 (55.6)	$\chi^2(1) = 3.02$, $p = 0.08$
Small for gestational age (SGA)[1] *n* (%)	17 (26.6)	0	
Bronchopulmonary dysplasia[2] *n* (%)	9 (14.1)		
Brain pathology in MRI[3] *n* (%)			
Normal	39 (60.9)		
Minor	12 (18.8)		
Major	13 (20.3)		
Maternal education at birth [4]*n* (%)			$\chi^2(1) = 3.46$, $p = 0.06$
≤ 12 years	17 (26.6)	27 (37.5)	
> 12 years	47 (73.4)	37 (51.4)	
FSIQ at five years of age, mean (SD) [min, max]	107.14 (15.83), [73, 140]	111.74 (12.94), [75, 136]	$t(134) = -1.85$, $p = 0.07$

1 Defined as a birth weight lower than −2 SD from age and gender specific Finnish growth charts.
2 Defined as the need for supplemental oxygen at 36 weeks of age.
3 For a detailed description of brain pathology, see Maunu et al. (2006, pp. 58–59).
4 Missing data from eight participants in the full-term sample.

maternal use of illicit drugs or alcohol during pregnancy. Table 1 presents the neonatal characteristics and maternal education information.

Written consent was obtained from all the participants and their parents, after they had received oral and written information regarding the study. The PIPARI study protocol, including the present study, was approved by the Ethics Review Committee of the Hospital District of Southwest Finland in December 2000 and January 2012.

Procedures

The participants were assessed individually in a separate room at their schools by a trained research assistant. The testing consisted of two 25–30 minute sets of measures of mathematical skills and domain general cognitive abilities and a 10-minute break including moving around in the school. The tasks were conducted in the following order: SFON task, standardized reading skills test (i.e. ALLU), SFOR task, mathematics motivation questionnaire, number line estimation task, relaxing break activity outside of the testing room, curriculum based mathematics achievement test (i.e. Lukilasse), Woodcock-Johnson III math fluency, rational number test, NEPSY-II inhibition, WMTB-C block recall, NEPSY-II visuomotor precision and WISC-IV digit span. One important factor in the measurement of SFON and SFOR was ensuring that the participants are not made aware of the mathematical nature of the tasks by external means (e.g. experimenter mentioning mathematics or counting), neither prior to nor during the testing. As participants are not explicitly guided towards the mathematical nature of the tasks, any focusing on numerosity or quantitative relations by the participants in the task can be considered to be initiated spontaneously by the participant. The WISC-IV full-scale intelligence quotient of the children born very preterm was assessed in a different session.

Measures of mathematical skills

SFON was assessed using a modified version of the Model task (Hannula and Lehtinen, 2005). The materials of the new task were three A4-sized pictures of a zebra, a leopard and a crocodile. Similar pictures were placed in front of the experimenter and the child. The child was told that in this task the experimenter would make her zebra into a model and then turn the model upside down on the table (so the child could not see it any longer). Then the child had to make his or her zebra look exactly like the model one. After introducing the task in this way, the experimenter said, "Now, watch carefully. I am making this zebra into a model." The experimenter draws 11 stripes starting from the rear of the zebra, shows the model to the child for five seconds and turns the model upside down. Then the experimenter says: "Now, make your zebra look like the model zebra." The procedure was repeated with a leopard with 13 drawn dots as the second item and a crocodile with 12 drawn teeth as the third item. Afterwards, the child was interviewed about what he or she paid attention to on the task. One point was given for each correctly drawn picture or for the mentioning of counting in the interview. The maximum score for the task was three points. Cronbach's alpha was 0.68.

SFOR was measured using the teleportation task (McMullen et al., 2016). On the first trial, the participant was asked to describe how two sets of objects (e.g. common food items) changed during a transformation (i.e. teleportation to another planet). The objects all changed color, shape, and number of objects in a uniform way (e.g., red objects become blue, milk becomes juice, the number of objects is divided by two). An example is shown in Fig. 1. On the second trial, the participant was shown a different amount of the original objects and was asked to draw what he or she expected to arrive based on what happened the previous trial. Then the two trials were repeated with new items. In total, there were four trials on the teleportation task lasting 1.5 minutes each. Points were given for each multiplicative relation mentioned or drawn, with a maximum of 3 points per trial and 12 points for the task. Cronbach's alpha was 0.74.

Guided focusing on quantitative relations was measured to determine the participants' ability to focus on and describe multiplicative relations imbedded in the SFOR tasks, when explicitly guided to do so, an important part of distinguishing spontaneous mathematical focusing tendencies from the underlying mathematical skills and knowledge (McMullen et al., 2019). After participants completed the SFOR task. Participants were presented with the first item from the teleportation task with explicit instructions to "describe how the items were divided" (McMullen et al., 2016). Students were given one point for correctly describing the multiplicative relation (e.g. "there were half as many items", "everything was multiplied by three").

Mathematics motivation was assessed with a modified Finnish version of the Fennema–Sherman Mathematics Attitude Scales (Metsämuuronen, 2009), which participants filled out themselves. This scale included 18 items with scoring from 1 to 5 points (i.e. disagree to agree). The scale measured four aspects of mathematical motivation: liking mathematics (5 items), self-concept in mathematics (5 items), utility in mathematics (5 items), and mathematics anxiety (3 items). Cronbach alpha for the total scale was 0.90.

Whole number line estimation was assessed with the number line estimation task (Siegler & Opfer, 2003). Participants placed values on a number line ranging from 0 to 1000. The participant was shown where the middle number is on the line. The participant needed to estimate 22 whole numbers. The mean absolute error (%) was calculated from the difference between the estimated and actual position of the number. Cronbach's alpha was 0.62 with the 22 items. Including fewer items would have improved the alpha, but not changed the results.

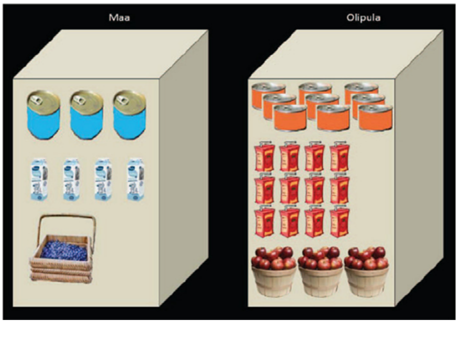

Fig. 1. Example of the teleportation task with the teleported material on the right. Picture with colours in online version.

Mathematics achievement was assessed with the 5th grade version of the Lukilasse (Häyrinen, Serenius-Sirve & Korkman, 1999), a curriculum-based standardized test. It consists of arithmetic problems, transcoding between verbal and Arabic digit numbers, converting problems between units of length and volume, and written problems. The test includes items with fractions and decimals. The maximum score is 21 points. Cronbach's alpha was 0.68.

Arithmetic fluency was measured using the math fluency subtest from Woodcock-Johnson III, which has been widely validated in previous studies (Woodcock et al., 2001). Participants were asked to complete as many single-digit basic arithmetic problems (addition, subtraction and multiplication) out of a two-page set of 160 as they can in three minutes. The number of correct answers was recorded, with a maximum of 160 points.

Rational number magnitude knowledge was assessed with 12 items adapted from Stafylidou and Vosniadou (2004), which included the comparison of fractions and/or decimals (e.g., "Circle the larger fraction. If the numbers are equal circle both: 5/8; 4/3") and the ordering of fractions or decimals (e.g., "Put the numbers in order from smallest to largest: 6/8; 2/2; 1/3"). A point was given for each correct answer with a maximum of 12 points. Cronbach's alpha was 0.70.

Measures of domain general cognitive abilities and reading skills

Naming speed, inhibition and shifting were assessed using the NEPSY-II inhibition subtest (Korkman, Kirk, & Kemp, 2008) with three parts assessing naming speed, response inhibition and shifting. In the task, the participant is shown a series of white and black circles and squares, and should either name the shapes, the opposite shape, or shift between alternating responses, as quickly as possible. The time for completion and errors were recorded. The time for completion in the naming task was used as a measure of naming speed. Standardized combined scores (i.e. combining errors and time), with a normative mean of 10, were used for inhibition and shifting. These are age-specific and standardized from the Finnish population.

Visuospatial processing was assessed with the block recall subtest from the WMTB-C (Pickering & Gathercole, 2001). In the task, the participant repeats sequences tapped by the administrator on identical wooden blocks, which are attached to a board in random positions. The total number of correct sequences was recorded.

Verbal working memory was measured using the backwards digit span subtest from the WISC-IV (Wechsler, 2003), which has been validated in the Finnish population. In the backwards condition, series of digits are read by the administrator and the participant repeats them in the reverse order. The total number of correct series was recorded.

Reading skills were measured with the word recognition subtest from ALLU (Ala-asteen Lukutesti; Lindeman, 2000), which is a standardized reading test for primary school. The participant had 3.5 minutes to read 78 chains of words, with two to four words in each chain, and separate the words with a line. One point was given for each correctly recognized word.

Data analysis

Data was analyzed using IBM SPSS Statistics 27. Chi-square test and independent samples *t*-test were used to compare groups on demographic variables. Independent samples *t*-test was used to examine group differences in mathematical skills, domain general cognitive abilities, and reading skills. The Holm-Bonferroni procedure (Holm, 1979) was used to account for familywise error rate by adjusting the alpha level to the number of dependent variables. This was calculated separately for measures of mathematical skills and domain general cognitive ability. Cohen's *d* was calculated for determining effect sizes. A chi-square test was used to compare the groups in guided focusing on quantitative relations. Univariate analyses of covariance (ANCOVA) were conducted to measure whether domain general cognitive abilities accounted for the group differences in mathematical skills. In addition, univariate ANCOVAs were conducted adjusting for gender and maternal education level, separately.

Prior to significance testing, all dependent variables were checked for normality of distributions and extreme values (i.e. scores > 3 SD from the median). One child born very preterm was excluded from the arithmetic fluency analysis because he or she performed 3 SD above the median of the preterm sample and 3 SD above the mean of the whole sample. This may have been a measurement error as the participant had average skills otherwise. Some dependent variables (e.g. SFON and SFOR) were not normally distributed due to skewness; consequently, additional non-parametric Mann-Whitney U tests were performed to verify the *t*-test result. All results were retained.

Results

Table 2 reports on the descriptive statistics for children born very preterm and full-term on the mathematical and cognitive measures. Table A1 in the appendix reports on the correlations between the mathematical measures and shows that all the measured mathematical skills are significantly related to mathematics achievement.

The full-term group had significantly higher performance than the preterm group on arithmetic fluency, SFON, and SFOR with small to medium effect sizes. Although the group difference in mathematics achievement was not statistically significant, 21.9% (14/64) of the preterm group scored below the 10th percentile, according to Finnish normative data, compared to 6.9% (5/72) of the full-term group, $\chi^2(1) = 6.28$, $p = 0.01$. Moreover, there were no significant group differences in whole number line estimation, rational number magnitude knowledge, mathematics motivation, or reading skills. In addition, the percentage of participants with a correct answer on the guided focusing on quantitative relations did not differ between the preterm and full-term groups, preterm: 88%, full-term: 90%, $\chi^2(1) = 0.27$, $p = 0.61$.

Group differences in mathematical skills after controlling for cognitive abilities

Next, we examined whether cognitive abilities account for the differences in mathematical skills between children born very preterm and full-term. The preterm group had a significantly slower naming speed and lower visuospatial processing than the full-term group (Table 2). Group differences were not statistically significant for inhibition, shifting, and verbal working memory. Univariate ANCOVAs were conducted to examine if the group differences in arithmetic fluency, SFON, and SFOR remained significant after controlling for naming speed and visuospatial processing. The group difference in arithmetic fluency was no longer statistically significant, $F(1,131) = 2.88$, $p = 0.09$, $\eta_p^2 = 0.02$, as naming speed explained a large amount of variance in arithmetic fluency, $F(1,131) = 22.11$, $p < 0.001$, $\eta_p^2 = 0.14$. However, the group differences remained significant for both SFON, $F(1,132) = 4.39$, $p = 0.04$, $\eta_p^2 = 0.03$, and SFOR, $F(1,132) = 4.62$, $p = 0.03$, $\eta_p^2 = 0.03$.

Additionally, within the children born very preterm, neonatal characteristics (gestational age, birth weight, SGA, brain pathology) were not significantly associated with the mathematical or cognitive measures. There were no group by gender interactions for the mathematical or cognitive measures, nor main effects of gender in the full-term group. Within the preterm group, boys had significantly lower performance on SFON ($M_{male} = 0.9$, SD = 1.1; $M_{female} = 1.6$, SD = 1.2; $t(62) = -2.44$, $p = 0.02$, $d = 0.62$), and higher performance on rational number knowledge ($M_{male} = 8.8$, SD = 2.9; $M_{female} = 7.2$, SD = 3.0; $t(62) = 2.08$, $p = 0.04$, $d = 0.53$). The group difference in SFON remained significant even after controlling for gender, $F(1,133) = 6.59$, $p = 0.01$, $\eta_p^2 = 0.05$. Maternal education measured at birth had data missing from eight full-term born children. After adjusting for maternal education in the remaining children and taking into account multiple comparisons, the results remained the same.

Retrospective comparison of mathematical skills and cognitive abilities at the age of five years

The current results on mathematical skills at the age of 11 years were surprising as they suggest that children born preterm without severe cognitive impairments do not have lower performance in many mathematical skills. The non-significant differences could be due to selection bias such as our preterm sample may have only consisted of children with exceptionally high mathematical skills

Table 2
Measures of mathematical and reading skills, and domain general cognitive ability at the age of 11 years.

	Very preterm (n = 64) Mean (SD)	Full-term (n = 72) Mean (SD)	Mean Difference (95% CI)	*t* (134)	*p*	Cohen's *d*
Mathematical measures						
SFON	1.16 (1.18)	1.72 (1.09)	−0.57 (-0.95 to −0.18)	−2.90	0.004[a]	0.50
SFOR	4.30 (3.60)	6.08 (3.91)	−1.79 (-3.07 to −0.51)	−2.76	0.007[a]	0.47
Arithmetic fluency[1]	63.13 (15.36)	71.07 (15.98)	−7.94 (-13.30 to −2.59)	−2.93	0.004[a]	0.51
Mathematics achievement	8.98 (4.26)	10.39 (4.09)	−1.40 (-2.82 to 0.01)	−1.96	0.052	0.34
Whole number line estimation	4.89 (1.48)	4.61 (1.33)	0.27 (-0.20 to 0.75)	1.13	0.261	0.19
Rational number magnitude	8.14 (2.97)	8.49 (3.05)	−0.35 (-1.37 to 0.68)	−0.67	0.506	0.11
Mathematics motivation	3.79 (0.72)	3.99 (0. 52)	−0.20 (-0.41 to 0.02)	−1.83	0.070	0.32
Reading skills	129.53 (34.51)	133.03 (34.45)	−3.50 (-15.21 to 8.22)	−0.59	0.556	0.10
Domain general cognitive measures						
Visuospatial processing	26.54 (3.21)	28.56 (3.84)	−2.01 (-3.22 to −0.80)	−3.29	0.001[a]	0.56
Naming speed	21.34 (3.28)	19.97 (2.76)	1.37 (0.33 to 2.41)	2.65	0.009[a]	0.45
Inhibition[2]	11.58 (1.88)	11.82 (1.93)	−0.24 (-0.89 to 0.41)	−0.74	0.463	0.13
Shifting[2]	11.67 (2.32)	12.21 (2.78)	−0.54 (-1.41 to 0.33)	−1.21	0.227	0.21
Verbal working memory	6.52 (1.44)	6.78 (1.50)	−0.26 (-0.76 to 0.24)	−1.04	0.302	0.18

[1] One child born very preterm was excluded from the analysis due to a score 3SD above the median.
[2] NEPSY-II standardized combined score for accuracy and speed.
[a] Significant after Holm-Bonferroni correction.

to begin with. As all the children in the current sample had completed measures of mathematical and cognitive abilities also at the age of five years, it was possible to investigate retrospectively whether differences in mathematical and cognitive abilities were present in early childhood. Hannula-Sormunen et al. (2017) found that mathematical skill profiles differed between children born very preterm and full-term at the age of five in a partially overlapping sample, but they did not directly compare the overall levels of early mathematical and cognitive abilities. We expected that group differences would exist already in early mathematical skills of the current sample based on previous studies from the PIPARI research project. In the present study, we therefore compared the children born very preterm and full-term on their early mathematical skills and NEPSY-II cognitive abilities measured at the age of five years (see Appendix B for a description of the testing procedures). Table 3 reports on the descriptive statistics and independent samples *t*-tests for these comparisons.

At the age of five years, the children born very preterm had significantly lower performance than the children born full-term on number sequence knowledge, cardinality skills, visuospatial memory, and phonological processing. These differences were of medium effect size. After controlling for multiple comparisons, there was no significant group difference in digit naming. In addition, there was no statistically significant group difference in SFON at the age of five when measured with numerosities within the subitizing range (i.e. less than three).

Additionally, there were no group by gender interactions for the mathematical or cognitive measures. However, a significant main effect of gender was found for digit naming, $F(1,133) = 12.40$, $p < 0.001$, $\eta_p^2 = 0.09$, with boys outperforming girls in both groups (preterm: $M_{male} = 4.6$, SD = 3.1 and $M_{female} = 2.6$, SD = 2.5; full-term: $M_{male} = 5.8$, SD = 3.4 and $M_{female} = 4.1$, SD = 2.9). For this reason, we controlled for gender, and the group difference in digit naming became significant, $F(1,133) = 6.83$, $p = 0.01$, $\eta_p^2 = 0.05$. Importantly, dropout analysis showed the children born very preterm or full-term who were present at the age of five years, but not in the 11-year-old sample, did not significantly differ in their early mathematical skills from the current sample. Only differences were that the children born very preterm lost to follow-up had lower visuospatial skills ($p < 0.001$) and FSIQ ($p < 0.001$) than those in the current sample. There were no significant differences in neonatal data.

Table 3
Descriptive statistics of mathematical skill and domain general cognitive ability at age five years by group.

	N	Very preterm Mean (SD)	N	Full-term Mean (SD)	Mean Difference (95% CI)	*t*	*p*	Cohen's *d*
Mathematical measures								
Cardinality skills	64	3.51 (1.79)	72	4.61 (1.89)	−1.10 (-1.73 to −0.48)	−3.48	<0.001[a]	0.60
Number sequence	64	16.64 (12.06)	72	23.49 (13.02)	−6.85 (-11.12 to −2.57)	−3.17	0.002[a]	0.54
Digit naming	64	3.77 (3.01)	69	4.84 (3.19)	−1.07 (-2.14 to −0.01)	−2.00	0.048	0.35
SFON	64	4.27 (3.52)	72	4.18 (3.58)	0.09 (-1.12 to 1.29)	0.14	0.889	0.02
Domain general cognitive measures								
Visuospatial memory[1]	62	8.55 (2.36)	71	10.04 (1.99)	−1.49 (-2.24 to −0.75)	−3.96	<0.001[a]	0.69
Phonological processing[1]	64	9.14 (2.20)	72	10.56 (2.41)	−1.41 (-2.20 to −0.63)	−3.56	<0.001[a]	0.61
Verbal working memory[1]	62	8.94 (3.69)	72	9.93 (2.43)	−1.00 (-2.05 to 0.06)	−1.81	0.073	0.32
Naming speed[1,2]	63	9.57 (2.98)	70	10.36 (2.58)	−0.79 (-1.74 to 0.17)	−1.63	0.106	0.28

[1] NEPSY-II standardized scores.
[2] Combined score for accuracy and speed.
[a] Significant after Holm-Bonferroni correction.

Discussion

The aim of the present study was to compare 11-year-old children born very preterm and full-term on a variety of 5th grade mathematical skills and investigate whether domain general cognitive abilities explain the group differences. Encouragingly, curriculum-based mathematics achievement did not differ, neither did whole number line estimation and rational number magnitude knowledge. However, the children born very preterm had lower arithmetic fluency, SFON, and SFOR compared to children born full-term. Although the group difference in arithmetic fluency was explained by naming speed, the group differences in SFON and SFOR remained significant even after controlling for domain general cognitive indicators. A retrospective analysis of the participating children at the age of five years (i.e. before primary school) showed that the children born very preterm had lower early mathematical and cognitive skills than their full term born peers. Despite having lower early mathematical skills, children born very preterm appear to reach peer equivalent skill levels in many of the measured mathematical skills by the age of 11 years. These results were mirrored in the children's mathematics motivation and reading skills, giving additional support that children born very preterm can catch up in academic skills even if they lag behind during early years.

Mathematical skills at the age of 11 years

The current results are in contrast to many previous studies on mathematics achievement, but not arithmetic fluency, in children born preterm (Aarnoudse-Moens et al., 2009; Twilhaar et al., 2018). However, most of these studies used only one standardized measure of mathematics achievement, while the current study included a set of specific mathematical measures and a mathematics achievement test. In addition, the mathematics achievement test was curriculum based, unlike in previous studies (e.g. Simms et al., 2015). Even though the difference in mathematics achievement was not significant on the group level, children born very preterm were three times more likely (22% vs 7%) to have a normative score below the 10th percentile. In addition, the children born very preterm had lower arithmetic fluency compared to the children born full-term. This suggests that the previously found difficulties in arithmetic fluency in early school years (Clark & Woodward, 2015; Pritchard et al., 2009; Short et al., 2003), even within this cohort (first grade; Korpipää et al., 2019), persist in children born very preterm. However, the weaker arithmetic fluency was explained by slower processing speed, supporting other studies in the finding that processing speed may underlie at least some mathematical challenges (Mulder et al., 2010; Rose et al., 2011).

The current results corroborate a previous finding on whole number line estimation being comparable between children born very preterm and full-term in primary school (Simms et al., 2015). Similarly, accuracy in whole number magnitude comparisons was previously found to be peer equivalent in primary school children born very preterm (Guarini et al., 2014, 2020; Simms et al., 2015). Even though studies have found group differences in 11-year-olds on whole number estimation, including number line estimation, these samples have been with children born extremely preterm (i.e. < 26 we eks; Johnson et al., 2009; Simms et al., 2013). The distinction is important as mathematical difficulties are found to increase with lower gestational age (Aarnoudse-Moens et al., 2009; Wolke et al., 2015). In line with this previous research on whole number magnitude knowledge, there was no substantial difference in rational number magnitude knowledge between the children born very preterm and full-term. This suggests that even with the increased cognitive demands of determining rational number magnitudes (Jordan et al., 2013), there do not appear to be persistent differences between children born very preterm and full-term on numerical magnitude processing.

While the performance of children born very preterm did not differ from the children born full-term on most mathematical tasks closely aligned with mathematics taught in the classroom, they differed in their SFON and SFOR. This indicates that children born very preterm at school age may be less likely to spontaneously focus on and utilize mathematical features in their everyday life (Hannula and Lehtinen, 2005). Importantly, these differences were not entirely explained by the group differences in visuospatial processing or naming speed. Nor were deficits in mathematical knowledge needed on the tasks the cause of these differences. For instance, the numerosities used within the SFON task (i.e. 11 to 13) were well within the range of enumeration skills of 11-year-old children. There was also no significant group difference on the guided focusing on quantitative relations task, which requires students to recognize and describe a multiplicative relation used in the SFOR task. Children born very preterm only displayed poorer performance on the mathematically implicit SFOR task. Thus, despite being able to focus on and describe the multiplicative relations when explicitly guided to do so, children born very preterm were less likely to do so spontaneously, without explicit guidance.

Even though the domain general cognitive abilities measured in this study did not explain the group differences in SFON and SFOR, lower SFON and SFOR could be an extension of general attentional problems in children born very preterm into the domain of mathematics. The SFON and SFOR tasks require the participant to process multiple aspects of potentially relevant stimuli. As the SFON task had a main effect of gender amongst children born very preterm, the results could support attention related problems in boys born preterm (Mulder, Pitchford, Hagger, & Marlow, 2009). Then again, gender did not have a significant effect on the SFOR task. Furthermore, previous findings suggest that children born very preterm catch up on attentional abilities, such as selective attention, by the age of 11 years (Mulder et al., 2009, 2010). As to potential difficulties with sustained attention, the children's attention was always ensured to be on the task by the experimenter. Consistent with previous findings (Edens & Potter, 2013; Nanu et al., 2018), the lack of attention to mathematical features was also not due to a group difference in mathematics motivation.

Interestingly, children born very preterm did not have lower SFON in tasks using only numbers in the subitizing range at the age of five years, yet they had lower SFON at the age of 11 years in a counting based task. Whether this discrepancy is due to differences in how subitizing and counting based enumeration are triggered by SFON, needs further investigation targeting directly at attentional, goal-directed processes needed when enumeration is used in action (Hannula, 2005). Nevertheless, the SFON result at the age of 11 years supports previous research on persistently lower performance in counting related mathematical skills (e.g. Clayton et al., 2021). Further research on the causes for lower spontaneous mathematical focusing tendency in children born very preterm could enrichen our understanding on individual differences in the use of existing mathematical knowledge in everyday environments.

Potential protective factors in mathematical development

The current sample is with 11-year-old children who do not have a severe cognitive impairment (i.e. all children have FSIQ > 70) and who have not been retained. The lack of severe cognitive impairment could be one reason for the apparently positive outcomes in the mathematical development of the children born very preterm. There were also no significant group differences between the 11-year-old children born very preterm and full-term in the executive function measures assessing inhibition, shifting, and verbal working memory. Verbal working memory did not significantly differ even at the age of five years in this sample. Thus, the executive function level of the children born very preterm in the present study aligns with the result that most of their classroom-related mathematical skills, including rational number magnitude knowledge, are peer equivalent at the age of 11 years. In contrast to previous studies (Litt et al., 2012; Simms et al., 2015), the lower visuospatial skills in children born very preterm, present at the ages of 5 and 11 years, did not appear to significantly hinder their learning of most classroom-related mathematical skills and knowledge.

In the PIPARI research project, children born very preterm generally have fewer problems with executive functions than in previous studies (Nyman et al., 2019b). The authors also noted that problems with working memory were related to lower FSIQ. Clayton et al. (2021) suggest that domain general cognitive abilities, such as weaker working memory, are the cause of persistent mathematical problems in children born very preterm. Our results are also in line with previous studies that have found lower mathematical skills in relation to weaker executive functioning in children born preterm and full-term (Bull & Scerif, 2001; Clark & Woodward, 2015; Rose et al., 2011; Simms et al., 2015, 2013; Twilhaar et al., 2020). Together with our results, these findings suggest that with the support of executive functions and lack of severe IQ deficits, lower early mathematical skills of children born very preterm can be overcome through formal schooling.

Educational implications

Overall, the present findings imply that 11-year-old children born very preterm without severe cognitive impairment do not differ from full-term in most of the currently measured classroom-related aspects of 5th grade mathematical knowledge or reading skills. These results are very encouraging because they show that children born very preterm can reach peer equivalent performance on many mathematical skills by 11 years of age, despite having significantly lower mathematical skills at the age of five years. While the children born very preterm had poorer speeded retrieval of arithmetic facts, this difference appeared to stem mainly from their slower processing speed. Likewise, Guarini et al. (2020) showed that 10-year-old children born very preterm compared to full-term had slower reaction times, but not accuracy, when comparing non-symbolic magnitudes. This implies difficulties with timed assessments due to slower retrieval of information rather than a lack of mathematical knowledge.

These results yield support to Nyman et al. (2019a) findings on peer equivalent teacher-ratings for mathematical skills at this age within the PIPARI research project. The authors suggested it could be due to the academic support services offered to Finnish students. It could also be due to the fewer executive function problems found within this cohort compared to previous literature (e.g. Nyman et al., 2019b). Moreover, the maternal education level amongst the children born very preterm in this sample may have been a protective factor (Nyman et al., 2019b; Doyle et al., 2015).

As mathematics is embedded in almost everything we do, the importance of mathematics is not only in the classroom. While the children born very preterm appeared to cope relatively well with classroom-based mathematical tasks, they may have a disadvantage in recognizing when mathematics is relevant in novel everyday-like situations. A wide variety of skills, ranging from the formal to informal, are necessary for becoming a mathematically literate member of society (Gravemeijer et al., 2017). The results of the present study suggest that children born very preterm could benefit from support with applying mathematical skills in everyday situations. This could support their arithmetic fluency and later mathematical skills.

Limitations and future directions

As the sample size of the present study is relatively small, the results should be interpreted with caution. Nevertheless, these results warrant more research into the causes for lower SFON and SFOR in children born very preterm and why the group differences in SFON were not present at the age of five years. In this study, we have only one or two tasks, with differing number ranges and task contexts, which could explain the lack of group differences at the age of five years. Future research should, first of all, measure SFON and SFOR using more tasks (McMullen et al., 2019). In addition, a battery of cognitive tasks would be needed to investigate more closely how executive functions and attentional abilities are related to SFON and SFOR in children born preterm. Finally, repeated measures analysis using a longitudinal design with similar tasks would be fruitful, as the developmental trajectories of SFON and SFOR tendency in children born preterm are not known. Likewise, longitudinal designs would be beneficial for discovering the developmental trajectories that allow children born very preterm to reach full-term equivalent performance in classroom-based mathematical skills.

The present study highlights the need to measure specific mathematical skills, when examining the effects of preterm birth, and continue follow-ups to examine long-term effects. Even though the children born very preterm had grade equivalent rational number magnitude knowledge, the rational number test included only a limited number of questions on the ordering and comparison of fractions and decimals. Consequently, the results cannot indicate whether they have difficulties with other aspects of rational number knowledge, nor their capability to apply this knowledge in other situations, such as everyday environments. Future research should examine this, as the adaptive use of mathematical skills is a requirement in many professions in the 21st century (ACME, 2011).

Acknowledgements

This study was supported by a grant from the Academy of Finland (278579) and (336068) to the last author and by the Foundation for Paediatric Research, Finnish Cultural Foundation, and Arvo and Lea Ylppö Foundation to Pipari Study Group. We warmly thank all the children, their parents and the schools for their participation. The PIPARI Study Group: Mikael Ekblad, MD, PhD; Satu Ekblad, RN; Eeva Ekholm, MD, PhD; Annika Eurola, MD; Linda Grönroos, MD; Leena Haataja, MD, PhD; Laura Haveri, BA; Minttu Helin, MD; Mira Huhtala, MD, PhD; Jere Jaakkola, BM; Eveliina Joensuu, MD; Max Karukivi, MD, PhD; Pentti Kero, MD, PhD; Riikka Korja, PhD; Katri Lahti, MD, PhD; Helena Lapinleimu, MD, PhD; Liisa Lehtonen, MD, PhD; Tuomo Lehtonen, MD; Marika Leppänen, MD, PhD; Annika Lind, PhD; Jaakko Matomäki, MSc; Jonna Maunu, MD, PhD; Petriina Munck, PhD; Laura Määttänen, MD; Anna Nyman, PhD; Riitta Parkkola, MD, PhD; Päivi Rautava, MD, PhD; Katriina Saarinen, Physiotherapist; Tiina Saarinen, MD; Susanna Salomäki, BA; Virva Saunavaara, PhD; Sirkku Setänen, MD, PhD; Matti Sillanpää, MD, PhD; Suvi Stolt, PhD; Karoliina Uusitalo, MD; Milla Ylijoki, MD, PhD. The authors have no interests that might be perceived as posing a conflict or bias.

Appendix A

See Table A1.

Table A1
Pearson correlations for the mathematical measures for the whole sample (n = 136).

	1	2	3	4	5	6
1. SFON	-					
2. SFOR	.22*	-				
3. Arithmetic fluency	.29***	.28**	-			
4. Mathematics achievement	.36***	.45***	.56***	-		
5. Whole number line estimation	-.09	-.25**	-.30***	-.38***	-	
6. Rational number magnitude	.15	.41***	.40***	.60***	-.37***	-
7. Mathematics motivation	.13	.19*	.37***	.42***	.-36***	.37***

Note. Mathematics achievement correlations (n = 135) due to exclusion of one participant born very preterm.
*$p < 0.05$ ** $p < 0.01$ ***$p < 0.001$

Appendix B

Retrospective analysis of the sample at the age of five years

1. Methods and procedures

The participants were assessed individually in a separate room by a psychologist at the University Hospital research unit at the age of five years (+0–2 months), more details from Hannula-Sormunen et al. (2017). There was no significant difference in the mean of days over five years of age between the children born very preterm and full-term, $M_{preterm}$ = 14.33, SD = 9.79, $M_{full\text{-}term}$ = 13.76, SD = 11.64, $t(134) = 0.30$, $p = 0.76$. The participants completed the following mathematical tasks and NEPSY-II subtests: SFON imitation tasks, "give a number" task, number sequence production, digit naming, phonological awareness, speeded naming, word list interference, and memory-for-design. There were short breaks between tasks when needed. The participants completed the WPPSI–R full-scale intelligence quotient in a different session.

1.1. Measures of mathematical skills

SFON was measured with two imitation tasks, the Parrot task and the Backpack task, modified from (Hannula and Lehtinen, 2005). The tasks included very small numerosities (i.e. one and two), which all children should recognize. The scoring was identical to Hannula-Sormunen et al. (2017): "The child was scored as focusing on numbers if she or he produced the correct numerosity and/or was observed presenting any of the mentioned quantifying acts." The maximum score for each task was four points with a total of eight points. Cronbach's alpha for the two tasks was 0.96.

In the parrot task, a blue toy parrot (capable of swallowing) was placed on the table in front of the participant and a plate of red glass berries was placed in front of the parrot. The experimenter first introduced the materials and then said, "Watch carefully what I do, and then you do just like I did". Then, the experimenter put two berries one at a time into the parrot's mouth, and these dropped with a bumping sound into the parrot's stomach. Afterwards, the participant was told, "Now you do exactly like I did". The following three trials consisted of one, two and one berries.

In the Backpack task, there was an empty blue bag and a basket with eight plastic, natural-sized oranges and eight pears. The tester sat opposite the participant, while holding the bag open on his or her lap. The basket was placed on the table next to them. The tester took two pears from the basket one at a time and put them into the backpack without letting the participant see inside the backpack while saying, "Let's play going outdoors and packing the backpack. Look carefully what I do. Then, you do it just like I did. Look, I do it. Now, please do it just like I did". After the participant had imitated the tester, no feedback was given, and the fruits were put back into the basket. In the second trial, the tester put one orange into the backpack. The next trials were carried out similarly but with a red backpack, and eight tomatoes and eight lemons in a different basket. The tester put two tomatoes (third trial) and one lemon (fourth trial) into the backpack.

Number sequence production was measured by asking the participant to count aloud from one onwards. The counting was stopped at 50. The participant had two trials and the highest number counted was recorded. The maximum score was 50.

Cardinality skills were assessed with the "give-a-number" task (Wynn, 1990). The participant was asked to place a number of items from a box onto the table. The numbers requested were 3, 5, 7, 9, 13, 19, and 23, respectively. The materials were seven different sets of painted wooden figures (flat). If the participant gave the wrong amount, the same trial was repeated. The task ended after two mistakes with the same number. A correct answer on the first trial was given one point and on the second trial was given 0.5 points, with a maximum of seven points.

Digit naming was measured with the digit naming subscale from the TEMA 3 (Ginsburg & Baroody, 2003). The participant named 15 different visually presented Arabic numerals, varying from one to four digits. One point was given for each correct answer with a maximum of 15 points. Cronbach's alpha was 0.87.

1.2. Measures of domain general cognitive abilities

The domain general cognitive abilities were measured with the Finnish NEPSY-II (Korkman, Kirk, & Kemp, 2008). Standardized norm scores (normative mean of 10) were used for all subtests.

Phonological processing was measured using the phonological awareness subtest. In the first part, the participant identifies spoken words from word segments. In the second part, the participant first repeats a word and then needs to omit or substitute a syllable or phoneme to form a new word.

Speeded naming is a measure of a participant's semantic access to over-learned items. In this subtest, the participant names arrays of figures, as quickly as possible, according to their color, shape and size. A combined score (i.e. combining speed and accuracy) was used.

Verbal working memory was assessed using the word list interference subtest. The participant listened to pairs of word series, progressing in length. For each pair, the participant first repeated each series of words immediately after it was presented. Finally, both series of words were repeated in the order of presentation. The total recall score was used.

Visuospatial memory was measured with memory-for-design subtest. The participant was shown a series of non-figure designs placed on a grid. After each presentation, the designs were removed from view. The participant needed to select the appropriate designs from a set of cards and place them in the same location on the empty grid. The participant was not penalized for placing too many cards but was instructed on how many cards should be placed.

References

Aarnoudse-Moens, C. S. H., Weisglas-Kuperus, N., Van Goudoever, J. B., & Oosterlaan, J. (2009). Meta-analysis of neurobehavioral outcomes in very preterm and/or very low birth weight children. *Pediatrics, 124*(2), 717–728. https://doi.org/10.1542/peds.2008-2816.

ACME (Advisory Committee on Mathematics Education). (2011). *Mathematical needs: Mathematics in the workplace and in higher education. Report of the ACME Mathematical Needs*. Retrieved from http://www.acme-uk.org/media/7624/acme_theme_a_final%20(2).pdf.

Alanko, O., Niemi, P., Munck, P., Matomäki, J., Turunen, T., Nurmi, J. E., ... Rautava, P., & the PIPARI Study Group. (2017). Reading and math abilities of Finnish school beginners born very preterm or with very low birth weight. *Learning and Individual Differences, 54*, 173–183. https://doi.org/10.1016/j.lindif.2017.01.022.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working Memory. *Developmental Neuropsychology, 19*(3), 273–293.

Clark, C. A. C., & Woodward, L. J. (2015). Relation of perinatal risk and early parenting to executive control at the transition to school. *Developmental Science, 18*(4), 525–542. https://doi.org/10.1111/desc.12232.

Clayton, S., Simms, V., Cragg, L., Gilmore, C., Marlow, N., Spong, R., & Johnson, S. (2021). Etiology of persistent mathematics difficulties from childhood to adolescence following very preterm birth. *Child Neuropsychology, 1–17*. https://doi.org/10.1080/09297049.2021.1955847.

DeWolf, M., Bassok, M., & Holyoak, K. J. (2015). From rational numbers to algebra: Separable contributions of decimal magnitude and relational understanding of fractions. *Journal of Experimental Child Psychology, 133*, 72–84. https://doi.org/10.1016/j.jecp.2015.01.013.

Doyle, L. W., Cheong, J. L. Y., Burnett, A., Roberts, G., Lee, K. J., & Anderson, P. J. (2015). Biological and social influences on outcomes of extreme-preterm/low-birth weight adolescents. *Pediatrics, 136*(6), e1513–e1520. https://doi.org/10.1542/peds.2015-2006.

Edens, K. M., & Potter, E. F. (2013). An exploratory look at the relationships among math skills, motivational factors and activity choice. *Early Childhood Education Journal, 41*(3), 235–243. https://doi.org/10.1007/s10643-012-0540-y.

Ginsburg, H., & Baroody, A. (2003). *Test of Early Mathematics Ability-Third Edition*. Austin, TX: Pro-Ed.

Gravemeijer, K., Stephan, M., Julie, C., Lin, F.-L., & Ohtani, M. (2017). What mathematics education may prepare students for the society of the future? *International Journal of Science and Mathematics Education, 15*, 105–123. https://doi.org/10.1007/s10763-017-9814-6.

Guarini, A., Sansavini, A., Fabbri, M., Alessandroni, R., Faldella, G., & Karmiloff-Smith, A. (2014). Basic numerical processes in very preterm children: A critical transition from preschool to school age. *Early Human Development, 90*(3), 103–111. https://doi.org/10.1016/j.earlhumdev.2013.11.003.

Guarini, A., Tobia, V., Bonifacci, P., Faldella, G., & Sansavini, A. (2020). Magnitude comparisons, number knowledge and calculation in verypreterm children and children with specific learning disability: A cross-population study using eye-tracking. *Journal of Learning Disabilities*. https://doi.org/10.1177/0022219420950651.

Hannula, M. M. (2005). *Spontaneous focusing on numerosity in the development of early mathematical skills*. Turku, Finland: Painosalama.

Hannula, M. M., & Lehtinen, E. (2005). Spontaneous focusing on numerosity and mathematical skills of young children. *Learning and Instruction, 15*(3), 237–256. https://doi.org/10.1016/j.learninstruc.2005.04.005.

Hannula, M. M., Lepola, J., & Lehtinen, E. (2010). Spontaneous focusing on numerosity as a domain-specific predictor of arithmetical skills. *Journal of Experimental Child Psychology, 107*(4), 394–406. https://doi.org/10.1016/j.jecp.2010.06.004.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*, 65–70.

Hannula-Sormunen, M. M., Nanu, C. E., Laakkonen, E., Munck, P., Kiuru, N., & Lehtonen, L. (2017). Early mathematical skill profiles of prematurely and full-term born children. *Learning and Individual Differences, 55*, 108–119. https://doi.org/10.1016/j.lindif.2017.03.004.

Häyrinen, T., Serenius-Sirve, S., & Korkman, M. (1999). *Lukilasse*. Helsinki, Finland: Psykologien Kustannus Oy.

Johnson, E. O., & Breslau, N. (2000). Increased risk of learning disabilities in low birth-weight boys at age 11 years. *Biological Psychiatry, 47*(6), 490–500.

Johnson, S., Hennessy, E., Smith, R., Trikic, R., Wolke, D., & Marlow, N. (2009). Academic attainment and special educational needs in extremely preterm children at 11 years of age: The EPICure study. *Archives of Disease in Childhood: Fetal and Neonatal Edition, 94*(4). https://doi.org/10.1136/adc.2008.152793.

Jordan, N. C., Hansen, N., Fuchs, L. S., Siegler, R. S., Gersten, R., & Micklos, D. (2013). Developmental predictors of fraction concepts and procedures. *Journal of Experimental Child Psychology, 116*(1), 45–58. https://doi.org/10.1016/j.jecp.2013.02.001.

Korkman, M., Kirk, U., & Kemp, S. L. (2008). *NEPSY II: Lasten neuropsykologien tutkimus [NEPSY-II: A developmental neuropsychological assessment]* (2nd ed.). Helsinki, Finland: Psykologien Kustannus Oy.

Korpela, S., Nyman, A., Munck, P., Ahtola, A., Matomäki, J., Korhonen, T., ... Haataja, L. (2018). Working memory in very-low-birthweight children at the age of 11 years. *Child Neuropsychology, 24*(3), 338–353. https://doi.org/10.1080/09297049.2016.1260101.

Korpipää, H., Niemi, P., Aunola, K., Koponen, T., Hannula-Sormunen, M., Stolt, S., ... Rautava, P., & the PIPARI Study Group. (2019). Prematurity and overlap between reading and arithmetic: The cognitive mechanisms behind the association. *Contemporary Educational Psychology, 56*, 171–179. https://doi.org/10.1016/j.cedpsych.2019.01.005.

Kucian, K., Kohn, J., Hannula-Sormunen, M., Richtmann, V., Grond, U., Käser, T., ... von Aster, M. (2012). Kinder mit Dyskalkulie fokussieren spontan weniger auf Anzahligkeit. *Lernen Und Lernstörungen, 1*(4), 241–253. https://doi.org/10.1024/2235-0977/a000024.

Lindeman, J. (2000). *Ala-asteen Lukutesti (ALLU) [Standardized, comprehensive school reading test]*. Jyväskylä, Finland: Gummerus.

Litt, J. S., Taylor, H. G., Margevicius, S., Schluchter, M., Andreias, L., & Hack, M. (2012). Academic achievement of adolescents born with extremely low birth weight. *Acta Paediatrica, 101*(12), 1240–1245. https://doi.org/10.1111/j.1651-2227.2012.02790.x.

Maunu, J., Kirjavainen, J., Korja, R., Parkkola, R., Rikalainen, H., Lapinleimu, H., & the PIPARI Study Group. (2006). Relation of prematurity and brain injury to crying behavior in infancy. *Pediatrics, 118*, 57–76.

McMullen, J., Chan, J. Y., Mazzocco, M. M. M., & Hannula-Sormunen, M. M. (2019). Spontaneous mathematical focusing tendencies in mathematical development and education. In A. Norton & M. W. Alibali (Eds.), *Constructing Number: Merging Perspectives from Psychology and Mathematics Education* (pp. 69–86). Springer. https://doi.org/10.1007/978-3-030-00491-0_4.

McMullen, J., Hannula-Sormunen, M. M., Laakkonen, E., & Lehtinen, E. (2016). Spontaneous focusing on quantitative relations as a predictor of the development of rational number conceptual knowledge. *Journal of Educational Psychology, 108*(6), 857–868. https://doi.org/10.1037/edu0000094.

McMullen, J., Hannula-Sormunen, M. M., & Lehtinen, E. (2014). Spontaneous focusing on quantitative relations in the development of children's fraction knowledge. *Cognition and Instruction, 32*(2), 198–218. https://doi.org/10.1080/07370008.2014.887085.

Metsämuuronen, J. (2009). *Methods assisting assessment: methodological solutions for the national assessments and follow-ups in the Finnish National Board of Education*. Oppimistulosten arviointi 1/2009. Opetushallitus. Helsinki: Yliopistopaino. [In Finnish.]

Mulder, H., Pitchford, N. J., Hagger, M. S., & Marlow, N. (2009). Development of executive function and attention in preterm children: A systematic review. *Developmental Neuropsychology, 34*(4), 393–421. https://doi.org/10.1080/87565640902964524.

Mulder, H., Pitchford, N. J., & Marlow, N. (2010). Processing speed and working memory underlie academic attainment in very preterm children. *Archives of Disease in Childhood: Fetal and Neonatal Edition, 95*(4), 267–272. https://doi.org/10.1136/adc.2009.167965.

Nanu, C. E., McMullen, J., Munck, P., Pipari Study Group & Hannula-Sormunen, M. M. (2018). Spontaneous focusing on numerosity in preschool as a predictor of mathematical skills and knowledge in the fifth grade. *Journal of Experimental Child Psychology, 169*, 42–58. https://doi.org/10.1016/j.jecp.2017.12.011.

Nyman, A., Korhonen, T., Lehtonen, L., & Haataja, L. (2019a). School performance is age-appropriate with support services in very preterm children at 11 years of age. *Acta Paediatrica*, 1–8. https://doi.org/10.1111/apa.14763.

Nyman, A., Korhonen, T., Munck, P., Parkkola, R., Lehtonen, L., & Haataja, L. (2017). Factors affecting the cognitive profile of 11-year-old children born very preterm. *Pediatric Research, 82*(2), 324–332. https://doi.org/10.1038/pr.2017.64.

Nyman, A., Munck, P., Koivisto, M., Hagelstam, C., Korhonen, T., Lehtonen, L., & Haataja, L. (2019b). Executive function profiles at home and at school in 11-year-old very low birth weight or very low gestational age children. *Journal of Developmental and Behavioral Pediatrics, 40*(7), 547–554. https://doi.org/10.1097/DBP.0000000000000689.

Pickering, S., & Gathercole, S. (2001). *Working memory test battery for children (WMTB-C). Manual*. London: Pearson Educational Ltd.

Pritchard, V. E., Clark, C. A. C., Liberty, K., Champion, P. R., Wilson, K., & Woodward, L. J. (2009). Early school-based learning difficulties in children born very preterm. *Early Human Development, 85*(4), 215–224. https://doi.org/10.1016/j.earlhumdev.2008.10.004.

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2011). Modeling a cascade of effects: The role of speed and executive functioning in preterm/full-term differences in academic achievement. *Developmental Science, 14*(5), 1161–1175. https://doi.org/10.1111/j.1467-7687.2011.01068.x.

Sansavini, A., Guarini, A., & Caselli, M. C. (2011). Preterm birth: Neuropsychological profiles and atypical developmental pathways. *Developmental Disabilities Research Reviews, 17*(2), 102–113. https://doi.org/10.1002/ddrr.1105.

Short, E. J., Klein, N. K., Lewis, B. A., Fulton, S., Eisengart, S., Kercsmar, C., . . ., & Singer, L. T. (2003). Cognitive and academic consequences of bronchopulmonary dysplasia and very low birth weight: 8-year-old outcomes. *Pediatrics, 112*(5), e359–1. https://doi.org/10.1542/peds.112.5.e359.

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., ... Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science, 23*, 691–697. https://doi.org/10.1177/0956797612440101.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14*, 237–243.

Simms, V., Gilmore, C., Cragg, L., Clayton, S., Marlow, N., & Johnson, S. (2015). Nature and origins of mathematics difficulties in very preterm children: A different etiology than developmental dyscalculia. *Pediatric Research, 77*(2), 389–395. https://doi.org/10.1038/pr.2014.184.

Simms, V., Gilmore, C., Cragg, L., Marlow, N., Wolke, D., & Johnson, S. (2013). Mathematics difficulties in extremely preterm children: Evidence of a specific deficit in basic mathematics processing. *Pediatric Research, 73*(2), 236–244. https://doi.org/10.1038/pr.2012.157.

Stafylidou, S., & Vosniadou, S. (2004). The development of students' understanding of the numerical value of fractions. *Learning and Instruction, 14*, 503–518.

Taylor, H. G., Espy, K. A., & Anderson, P. J. (2009). Mathematics deficiencies in children with very low birth weight or very preterm birth. *Developmental Disabilities Research Reviews, 15*(1), 52–59. https://doi.org/10.1002/ddrr.51.

Twilhaar, E. S., de Kieviet, J. F., Aarnoudse-Moens, C. S. H., van Elburg, R. M., & Oosterlaan, J. (2018). Academic performance of children born preterm: A meta-analysis and meta-regression. *Archives of Disease in Childhood - Fetal and Neonatal Edition, 103*(4), F322–F330. https://doi.org/10.1136/archdischild-2017-312916.

Twilhaar, E. S., De Kieviet, J. F., Van Elburg, R. M., & Oosterlaan, J. (2020). Neurocognitive processes underlying academic difficulties in very preterm born adolescents. *Child Neuropsychology, 26*(2), 274–287. https://doi.org/10.1080/09297049.2019.1639652.

Van Hoof, J., Degrande, T., McMullen, J., Hannula-Sormunen, M., Lehtinen, E., Verschaffel, L., & Van Dooren, W. (2016). The relation between learners' spontaneous focusing on quantitative relations and their rational number knowledge. *Studia Psychologica, 58*(2), 156–170 https://doi.org/10.21909/sp.2016.02.714.

Van Hoof, J., Janssen, R., Verschaffel, L., & Van Dooren, W. (2015). Inhibiting natural knowledge in fourth graders: Towards a comprehensive test instrument. *ZDM, 47*(5), 849–857. https://doi.org/10.1007/s11858-014-0650-7.

Verschaffel, L., Rathé, S., Wijns, N., Degrande, T., Van Dooren, W., De Smedt, B., & Torbeyns, J. (2018). Young children's early mathematical competencies: The role of mathematical focusing tendencies. In I. Erfjord (Ed.), *Mathematics education in the early years. Results from the POEM4 conference, 2018*. New York: Springer.

Wehrle, F. M., Kaufmann, L., Benz, L. D., Huber, R., O'Gorman, R. L., Latal, B., & Hagmann, C. F. (2016). Very preterm adolescents show impaired performance with increasing demands in executive function tasks. *Early Human Development, 92*, 37–43. https://doi.org/10.1016/j.earlhumdev.2015.10.021.

Wechsler, D. (2003). *WISC-IV, technical and interpretive manual*. London: Pearson Assessment.

Wolke, D., Strauss, V. Y. C., Johnson, S., Gilmore, C., Marlow, N., & Jaekel, J. (2015). Universal gestational age effects on cognitive and basic mathematic processing: 2 cohorts in 2 countries. *Journal of Pediatrics, 166*(6), 1410–1416.e2. https://doi.org/10.1016/j.jpeds.2015.02.065.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside.

Wynn, K. (1990). Children's understanding of counting. *Cognition, 36*, 144–193.