



The relationship between number line estimation and mathematical reasoning: a quantile regression approach

Carola Ruiz¹ · Saskia Kohnen² · Rebecca Bull¹

Received: 12 September 2022 / Revised: 8 May 2023 / Accepted: 9 May 2023 /
Published online: 23 May 2023
© The Author(s) 2023

Abstract

Number line estimation has been found to be strongly related to mathematical reasoning concurrently and longitudinally. However, the relationship between number line estimation and mathematical reasoning might differ according to children's level of performance. This study investigates whether findings from previous studies that show number line estimation significantly predicting mathematical reasoning replicate, and whether this relationship holds across several points of the mathematical reasoning distribution. Participants include 324 Singaporean children (162 girls, $M_{\text{age}}=74.1$ months, $SD_{\text{age}}=4.0$) in their second year of kindergarten who were assessed on the number line estimation task (0–10 and 0–100) and mathematical reasoning skills. The results replicate previous findings showing that higher accuracy on the number line estimation task is predictive of higher mathematical reasoning, for both the 0–10 and 0–100 number line. Quantile regressions show that performance on the number line task similarly predict mathematical reasoning across the performance distribution. Possible differences between the 0–10 and 0–100 number line's predictive capacity are discussed.

Keywords Quantile regression · Number line · Numeracy · Mathematics · Mathematical reasoning

Introduction

The use of number lines as a representation of numbers has been recommended as an effective tool for students to engage and develop mathematical ideas (Education Endowment Foundation, 2017; 2020). There is strong evidence for its use with low achieving students, as number lines allow for the representation of all real numbers, supporting the acquisition of basic to advanced mathematical concepts (Fuchs et al., 2021). Number lines can also support

✉ Carola Ruiz
carola.ruizhornblas@hdr.mq.edu.au

¹ School of Education, Macquarie University, Sydney, Australia

² School of Psychological Sciences, Macquarie University, Sydney, Australia

the acquisition of mathematical operations and are used in the context of temperature, time, graphs, etc. In the last few decades, numerous studies within the fields of psychology and education have tapped children's mental representation of numbers using the number line estimation task. Performance on the number line task has been associated with mathematical reasoning concurrently and longitudinally (Ellis et al., 2021). However, there is increasing evidence that the association between some mathematical skills and mathematical reasoning might differ based on children's mathematical reasoning ability (Fuchs et al., 2020; Hajovsky et al., 2020). These studies have explored the effects of the association of more complex mathematical skills (e.g., calculation fluency) with mathematical reasoning, and have found differing effects. While there is evidence that supports a differential association between some basic numerical skills (e.g., approximate number system, counting, number relations) and performance in more complex mathematical tasks (Bartelet et al., 2014; Devlin et al., 2022; Purpura & Logan, 2015), this has not been studied with mathematical reasoning as an outcome. Thus, it seems possible that a more basic numerical skill, such as number line estimation, could have a differential relationship with a complex mathematical skill such as mathematical reasoning. Understanding individual differences in this association is relevant for several reasons: i) it can aid in clarifying the role of number line estimation in mathematical reasoning, ii) it can help better understand the processes involved in mathematical reasoning, iii) it can guide the targeted use of educational interventions. The latter can help target interventions, which safeguards the appropriate allocation of limited economic and human resources, by maximizing the cost–benefit relationship of educational interventions. Considering that curricula worldwide include mathematical reasoning as a main component of mathematics learning (Australian Curriculum, Assessment and Reporting Authority, 2018; Department for Education, 2014; Ministry of Education, 2021; National Council for Teachers of Mathematics, 2000), this study aims to determine if accuracy on the number line task (0–10 and 0–100) differently predicts concurrent mathematical reasoning across the spectrum of mathematical reasoning performance.

Mathematical reasoning

Mathematical reasoning is emphasized as being a critical feature of mathematics learning (Herbert & Williams, 2023; Lindmeier et al., 2018; Olsson & Granberg, 2022), and it is a key concept in curricula worldwide (Australian Curriculum, Assessment and Reporting Authority, 2018; Department for Education, 2014; Ministry of Education, 2021; National Council for Teachers of Mathematics, 2000). This construct has been defined by the OECD (2018) as involving “evaluating situations, selecting strategies, drawing logical conclusions, developing and describing solutions, and recognizing how those solutions can be applied” (p.14). A variety of frameworks have been proposed (e.g., Ball & Bass, 2003; Jeannotte & Kieran, 2017; Lithner, 2008; Yackel & Hanna, 2003), highlighting the variability of interpretations of this concept. Lindmeier et al. (2018) suggests that mathematical reasoning, at its most basic, can be considered as “reasoning in mathematical situations” (p. 315). The authors propose that even though mathematical reasoning is usually conceptualized as requiring the use of complex knowledge and processes, there is still a form of mathematical reasoning that can take place before children are formally taught mathematics. An example of this is provided by Sumpter (2014) who analyses examples of individual and collective mathematical reasoning in the preschool years. Though the term mathematical reasoning is conceptualised in many different tasks, it is usually studied in the context of problem solving. This also applies to the current study. Children had to recognize the total number of items that result from combining

two sets, or from taking away items from a set, understand information presented in graphical format to answer questions, among others. Mathematical reasoning has been shown to longitudinally predict performance on high-stakes mathematics assessments (Nunes et al., 2007), above and beyond arithmetic skills (Nunes et al., 2012).

The number line estimation task

Number line estimation is usually operationalized by number line tasks. Several variations of this task have been proposed. In the research literature, and in educational settings, the most frequently used version of the task is the ‘number to position bounded’ task with Arabic numbers. In this task, a horizontal line is presented with labeled endpoints (bounded) which indicate the range of the number line (e.g., 0–10, 0–100, 0–1000). Participants are presented with a series of Arabic numbers which they place on the line based on their value (number to position). Differences have been found between versions of the task regarding what they index (Reinert et al., 2019) and how they relate to mathematical achievement (Schneider et al., 2018). The focus of this study is on the number to position, bounded number line task with Arabic numbers as it is the most frequently used version of the task.

Various explanations regarding what underlies performance on the number line estimation task have been put forward (see Dackermann et al., 2015; Yuan et al., 2020; for an extensive overview). An analysis of these theoretical explanations is beyond the scope of this study. However, considering what the number line task represents is important to interpret its relationship to other mathematical skills. Initial proposals suggested that the number line estimation task mainly indexes the mental representation of magnitude (Siegler et al., 2011). According to this account, performance on the number line reflects the way numbers are represented mentally. An alternative account advocates that performance on the bounded number line task indexes the use of proportion judgment skills (Cohen & Blanc-Goldhammer, 2011). Based on this account, performance on the number line reflects the strategies used to relate the number to be estimated with the numerical range of the number line. Other accounts have proposed ordinal processing (Podwysocki et al., 2019), familiarity with numbers (Ebersbach et al., 2008, 2015), place value understanding (Moeller et al., 2009) and spatial mapping (Sella et al., 2017) as underlying performance on the number line estimation task. Regardless of what underlies performance on the number line task, its relevance and relationship with other mathematical skills do not seem to be debated.

Number line estimation and its relationship with mathematical reasoning

Several studies have focused on the concurrent and longitudinal relationship between performance on the number line task and other mathematical skills across the lifespan. A meta-analysis conducted by Schneider et al., (2018) found an overall moderate correlation ($r=0.443$) between number line estimation and mathematical achievement (i.e., combination of all mathematical outcomes included in the meta-analysis). This association was shown to differ by age, and to be stronger for later than concurrent mathematical achievement. Neither the numerical range of the number line, nor the type of measure used to index mathematical achievement were significant moderators of the relationship between number line estimation and mathematical achievement. However, this study did not specifically differentiate mathematical reasoning as an outcome measure.

Ellis et al., (2021) replicated this study based on seven samples, differentiating mathematical reasoning as an outcome measure. The authors found an overall relationship between number line estimation and mathematical reasoning ($r=-0.36$), which was moderated by age and number range. Additionally, Zhu et al., (2017) found that performance on the 0–100 number line task predicted 12% ($\beta=-0.370$) and 9% ($\beta=-0.318$) of unique variance of mathematical reasoning, in grades 2 and 4, respectively, after accounting for age, non-verbal reasoning, attention and working memory. Similarly, Ouyang et al., (2021) found that performance on a similar number line task (0–100) during the second preschool year was predictive of mathematical reasoning in the third year of preschool ($\beta=-0.250$), even after accounting for the effect of several covariates ($\beta=-0.275$) (but see Clarke et al., 2020 for different results). These studies provide evidence of a relationship between number line estimation and mathematical reasoning. However, it is relevant to further analyse if and how the number range on the number line task moderates its relationship with mathematical reasoning. Importantly, none of these studies have examined whether the relationship between number line estimation and mathematical reasoning might differ depending on children's mathematical reasoning ability, even though previous evidence suggests this is the case for other early developing numerical abilities.

The effects of number range on the relationship between number line estimation and mathematical reasoning

Ellis et al., (2021) found that the relationship between number line and mathematical reasoning held across different number ranges (i.e., $r=-0.40$, $r=-0.51$, $r=-0.55$ for the 0–20, 0–100 and 0–1000 number ranges, respectively), though they found that the relationship with mathematical reasoning was significantly higher for the 0–1000 compared to the 0–20 number range. The authors explain that these findings could arise from limitations of the included samples that led to confounds between number range and presentation medium, as within the included studies different number ranges were presented in different formats (i.e., 0–20 in paper and pencil, while 0–100 and 0–1000 on a computer). However, they also suggest that if some number ranges become too simple for children to estimate, this might decrease the relationship between performance on the number line estimation task and mathematical reasoning. Additionally, as proposed by Moeller et al., (2009) and Ebersbach et al., (2008, 2015), children might engage differently with the number line estimation task depending on its number range. This could lead to a difference in the mathematical skills being used to perform the number line task. It is possible that children might rely on different mathematical skills when dealing with highly known as compared to lesser-known number ranges. For example, children may use ordinality skills to solve the number line task with a highly known number range (e.g., 0–10), while more complex skills such as place value understanding or proportional reasoning might be needed to accurately solve the number line task with lesser-known number ranges (e.g., 0–100). However, it is also possible that other differences might depend on the characteristics of the number range itself. For example, place value understanding might become more relevant within the 0–100 or 0–1000 number ranges, as compared to the 0–10 number range. Finally, it is also possible that as children's mathematical skillset diversifies, the strategies they choose to use to solve the number line task change. Though most of the explanations proposed regarding the skills underlying performance on the number line estimation task somewhat account for changes in proficiency throughout time (i.e., level of linearity of the mental representation of magnitude, number of benchmarks used for estimation), limited attention

has been paid to the possibility that different number ranges in the number line estimation task might be reflecting different underlying mathematical skills, and that these differences might be further impacted by a variety of factors such as age, familiarity with numbers, and other mathematical skills.

The effects of the level of mathematical reasoning on the relationship between number line estimation and mathematical reasoning

Previous findings indicate that the predictive capacity of some basic numerical skills to explain mathematical skills might vary across differing levels of performance. For example, Bartelet et al., (2014) found that counting only significantly predicted fact retrieval at the 84th percentile, though differences between the 16th, 50th and 84th quantiles were not statistically significant. Devlin et al., (2022) studied the capacity of different subdomains of early numeracy (i.e., numbers, number relations and number operations) in pre-kindergarten, kindergarten and first grade to predict mathematical achievement (i.e., an assessment battery tapping several mathematical skills) a year later. They found that number operations were significantly predictive of mathematical achievement at the 50th and 80th quantiles, but not the 20th quantiles for all three grades. Number relations significantly predicted mathematical achievement across all quantiles assessed for the pre-kindergarten cohort, but it was only significantly predictive at the 20th quantile for the kindergarten and first grade cohorts. Finally, numbers significantly predicted mathematical achievement at the 20th and 50th quantiles in pre-kindergarten, at the 50th and 80th quantiles in kindergarten and only at the 50th quantile in first grade. This highlights how different subdomains of early numeracy might differently impact mathematical achievement depending on the children's level of mathematical achievement, but also on their age. However, statistical comparisons across quantiles are not reported. Furthermore, Li et al., (2018) found that approximate number system performance in fourth grade significantly predicted concurrent strategy choice on a double-digit multiplication task for median and high quantiles (50th, 75th and 85th), but not for lower performance (15th and 25th), though statistical comparison between quantiles were not reported. Li et al., (2018) offer several possible interpretations for their findings. They initially suggest that children with poorer strategy choice also show poorer approximate number system performance; this would mean that they have to rely on more general cognitive skills such as executive functions as a compensatory mechanism. Additionally, the authors suggest that this could also be interpreted from the mutual inclusion/exclusion theory. Within this framework they suggest that by mutual inclusion, children with high performance in strategy choice also exhibit high approximate number system performance. High approximate number system performance together with high performance on the central executive (as reflected by their findings) would lead to high levels of problem solving. However, it is unclear based on these interpretations how strategy choice is impacted by the approximate number system. Relatedly, Purpura & Logan, (2015) found that approximate number system performance was a significant predictor of future early math competence for the 25th, but not the 50th or 75th percentiles when concurrent early math competence and math vocabulary were accounted for. However, comparisons across quantiles are not reported. The authors suggest that these findings are expected given that lower performance on early math competence would involve solving items with small quantities, for which reliance on the approximate number system would be higher than more advanced items which include larger quantities and making connections between them.

This phenomenon has also been observed in the predictive capacity of some mathematical skills to explain performance on mathematical reasoning. For example, Fuchs et al., (2020) with a sample of second graders found that even though calculation fluency was a significant predictor of mathematical reasoning across the performance distribution, its predictive strength was higher for the 80th compared to the 20th and 50th percentile (with there being no significant difference between the 20th and 50th percentiles). The authors propose two possible explanations for these findings. On the one hand, they suggest that children who exhibited weaker mathematical reasoning also probably have limited calculation skills, which could explain a weaker relationship between both skills for the lower and intermediate quantiles as compared to the higher quantile. This explanation resembles the one proposed by Li et al., (2018) in reference to the approximate number system performance not being predictive of strategy choice in multiplication for the lower quantiles. Alternatively, Fuchs et al., (2020) propose that children who showed low and intermediate mathematical reasoning performance tend to approach word problems by adding the numbers presented on the problems, without engaging with mathematical reasoning. This would lead to low accuracy on measures of mathematical reasoning (which in this case include subtraction problems for half of the items), while calculation accuracy might remain high. However, it is relevant to consider that, though not included in the comparison across quantiles, the estimate for calculation fluency was the lowest for the 90th percentile. Without further explanations, this would not be consistent with the authors' interpretations of their findings. Relatedly, Hajovsky et al., (2020) found that within a sample of kindergarten through grade 2 children, calculation fluency was negatively related to concurrent mathematical reasoning at the 10th percentile, but positively related at the 50th and 90th percentiles. The authors do not specifically interpret these findings but highlight the implications of accounting for non-linear relationships between cognitive and academic skills when having to prioritise which skills to include in the context of neuropsychological assessments with children. Finally, a similar phenomenon has been shown for the relationship between general cognitive processes and mathematical skills (Dong et al., 2020; Hajovsky et al., 2020), as well as mathematics self-concept and mathematical skills (Susperreguy et al., 2018). To the best of our knowledge, this has not been studied for the relationship between number line estimation and mathematical reasoning.

Performance on the number line estimation task could have a different role across differing levels of mathematical reasoning. The reasons behind this difference and its interpretations would depend on what performance on the number line represents. For example, under the assumption that performance on the number line estimation task mainly reflects the mental representation of magnitudes (Siegler et al., 2011), then one could speculate that the accuracy of the mental representation of magnitudes would possibly have a stronger effect for the initial acquisition of mathematical reasoning skills, but that once a certain minimum level of performance has been acquired, representations of magnitude would become less important. This would entail a stronger relationship between number line estimation and mathematical reasoning for lower than higher levels of mathematical reasoning and would be consistent with previous findings showing that performance of the approximate number system is significantly predictive of early math competence only for lower levels of performance (Purpura & Logan, 2015).

Alternatively, under the assumption that completing the number line task is a proportional judgment task (Cohen & Blanc-Goldhammer, 2011), therefore entailing proportion understanding and the ability to spatially reflect these proportions, then one could have different expectations for the relationship between the number line task and mathematical reasoning. In this case, understanding of proportions between numbers could reflect complex numerical knowledge, and would therefore only become necessary when solving problems

that require this complex knowledge. If this were the case, one could expect performance on the number line task to be more strongly related to mathematical reasoning for higher than lower ability levels. This would resemble Fuchs et al., (2020) findings in which calculation fluency was more strongly related to mathematical reasoning in the 80th percentile as compared to the 20th or 50th percentile (though this relationship was the lowest for the 90th percentile). Other interpretations would ensue when considering place value understanding, familiarity with numbers, ordinality or spatial mapping as underlying performance on the number line estimation task.

In summary, previous studies support a relationship between the number line estimation task and mathematical reasoning. However, determining if, and how, the relationship between performance on the number line task and mathematical reasoning differs across levels of performance is key. This could allow us to better understand the role that number line estimation plays in the development of mathematical reasoning, as well as to create differentiated and targeted educational interventions if required.

The present study

This study has two aims. The first aim is to replicate the findings from previous studies that show that higher accuracy on the number line estimation task is significantly associated with higher levels of mathematical reasoning. We use linear regression to fulfill this aim. Linear regression assumes homogeneity of the association between number line estimation and mathematical reasoning. However, it is possible that there is heterogeneity in whether and how the number line estimation task predicts mathematical reasoning. Therefore, the second aim is exploratory. We want to determine if the relationship between number line estimation and mathematical reasoning holds across several points of the performance distribution. We use a conditional quantile regression approach to fulfill this aim. This approach has the benefit that it estimates the relationship between the predictors of interest (0–10 and 0–100 number line) and the outcome (mathematical reasoning) at different points of the outcome's distribution (Petscher & Logan, 2014).

Materials and methods

Participants

The initial sample was composed of 347 children in the second year of kindergarten in Singapore (year before starting primary school). This sample represents all children who participated in a cross-sectional study, a smaller cohort of which was followed up as part of a longitudinal study reported elsewhere (Bull et al., 2021). Participants attended a total of 21 preschools which were selected from the government-supported, non-profit, and for-profit providers that are typical of the early childhood service options available to the local population. Due to children's absence from preschool on testing days, data for one or more variables was missing for a small number of participants. Therefore, data was analyzed for missingness. Little's (1988) Missing Completely at Random (MCAR) test was not significant, indicating that data was likely missing completely at random ($\chi^2 = 29.986$, $df = 20$, $p = 0.166$). As a result, listwise deletion was used and 23 observations were deleted from the dataset. The final sample comprised of 324 participants (162 girls, $M_{\text{age}} = 74.1$ months, $SD_{\text{age}} = 4.0$).

According to parents' reports, the sample consisted of 63.0% Chinese children, 21.0% Malay children, 9.3% Indian children and 5.2% children from other ethnicities. Ethnicity for five children was not reported. Comparison with census data indicated this sample had an under-representation of Chinese children (census = 74.2%) and an over-representation of Malay children (census = 13.3%; Department of Statistics Singapore, 2013). Information regarding parent's highest educational level achieved was reported for two hundred and 28 children – 8.3% did not have high school qualifications, 17.3% had high school qualifications, 22.2% had post-secondary qualifications and 22.6% had a university undergraduate or postgraduate qualification. Census data reports 31.2% without high school qualification, 18.8% high school qualification, 22.7% post-secondary qualification and 27.3% university qualification (Department of Statistics Singapore, 2013).

Procedure

The study received ethical approval by Nanyang Technological University. Prior to children's participation, parents provided written consent. All children were tested in English, which is the language of instruction, within their preschools and given a small gift for their participation. Testing time was approximately 1.5 h divided into two to three sessions, conducted on different days. Task order was randomized, with the restriction that both number line tasks had to be completed on different days.

Measures

Vocabulary and non-verbal reasoning skills were assessed to control for general cognitive functioning. Non-verbal reasoning was assessed through Raven's Progressive Matrices test and vocabulary was assessed through the Peabody Picture Vocabulary Test. In both cases, assessments were conducted following standard procedures. Gender was classified as male/female according to parent reports.

General cognitive skills

Vocabulary The Peabody Picture Vocabulary Test III (Dunn & Dunn, 1997) assesses children's receptive vocabulary. After listening to a word, the child selects one out of four images that best represent the meaning of the word they heard. The assessment was conducted following standard procedures. Accuracy is indexed as the number of correct responses.

Non-verbal reasoning Raven's Standard Progressive Matrices Test (Raven et al., 2003) assesses children's non-verbal reasoning. Children are presented with a matrix which is missing a section and they must choose which of the options provided corresponds to the missing section. The assessment was conducted following standard procedures. Accuracy is indexed as the number of correct responses.

Mathematical skills

Number Line Estimation (0–10) Children were presented with a number between one and nine, and a horizontal line with zero at the left and ten at the right. Children were asked to determine where the number should go on the horizontal line. Children initially completed one practice trial (five) to confirm they had understood the task instructions and were then assessed with all numbers between one and nine. Each number was presented twice (Siegler & Booth, 2004), and an average position was calculated. Children received no accuracy-based feedback throughout the task (including practice trials). Accuracy was indexed by percent absolute error $[(\text{number estimated} - \text{target number}) / \text{scale of number line}] \times 100$. Spearman-Brown split half reliability coefficient for this sample was 0.950. This number range was chosen as it was aligned with curriculum expectations (Ministry of Education, 2013), and would therefore represent a well-known number range.

Number Line Estimation (0–100) Children were presented with a number between one and 99, and a horizontal line with zero at the left and 100 at the right. Children were asked to determine where the number should go in the horizontal line. Children initially completed one practice trial (fifty) to confirm that they had understood the task and were then assessed with an additional 26 numbers (three, four, six, eight, 12, 14, 17, 18, 21, 24, 25, 29, 33, 39, 42, 48, 52, 57, 61, 64, 72, 79, 81, 84, 90, and 96). Numbers on the lower end of the number line were oversampled following Siegler and Booth (2004). Children received no accuracy-based feedback throughout the task (including practice trials). Accuracy was indexed by percent absolute error. Spearman-Brown split half reliability coefficient for this sample was 0.917. This number range was chosen to extend beyond well-known number ranges and curriculum expectations (Ministry of Education, 2013).

Mathematical Reasoning The Math Reasoning subtest from the Wechsler Individual Achievement Test II (WIAT-II, Wechsler, 2005) assesses the application of numerical and mathematical concepts in practical situations and the ability to solve verbally presented word problems aided by visual cues. Children might be required to answer items such as: “Abdul had three marbles in one hand and two marbles in the other. How many marbles did Abdul have in total?” which in some cases would be accompanied by pictures showing the marbles in Abdul’s hands. Other items might involve being able to answer questions based on information provided by a calendar or a ruler. Following standard procedures, testing was discontinued after six consecutive errors. Accuracy is indexed by number of correct responses.

As part of the original study the following measures were also assessed: Bracken School Readiness Test (colours, letters, numbers, sizes and shapes), non-symbolic comparison, non-symbolic to symbolic estimation, Numerical operations subtest from WIAT II and, parent questionnaire regarding expectations for learning. We had initially included performance on the Numerical Operations subtest as a dependent variable in this study, but later concluded that given the characteristics of the task (i.e., first items assessing number identification and counting, and following items assessing calculation) differences across quantiles in this task were more likely to reflect differences between skills than differences in performance across the same skill. Therefore, those analyses are not included in the study. None of the other variables were included as they were not aligned with the aims of the study.

Results

To address the aims of this study, we conducted two separate linear and quantile regressions with mathematical reasoning as dependent variable, one for each number range (i.e., 0–10, 0–100) as predictor.¹ All analyses were conducted in R (R Core Team, 2019). The linear regression was aimed at establishing if, as in previous studies, there was a mean association between performance on the number line task and mathematical reasoning. The Stats package in R was used to conduct these analyses. Including the quantile regressions had the purpose of determining if this association varied across the distribution of mathematical reasoning. The conditional quantile regression approach applied in this study allows to use the data from the whole sample to analyze the relationship between the predictors and outcome variables at different points of the outcome's distribution (Petscher & Logan, 2014). This approach prevents the loss of statistical power that would arise if independent linear regressions were conducted with separate subgroups based on their level of performance. Wenz (2019) points out that the quantile regression approach analyses the effects of the predictors in the 'conditional' distribution of the outcome variable, which is conditional on all the predictors included. Even though estimate coefficients from quantile regression are interpreted similarly to those of linear regression, this effect is not on the outcome variable itself, but on the conditional distribution of that variable. As more predictors are included in the regression, it is more likely that the conditional distribution of the outcome will differ from the unconditional distribution. We focused on the 10th, 25th, 50th, 75th and 90th percentiles for these analyses. We specifically chose the 10th and 25th percentiles as these are frequently used in the research literature as upper benchmarks to indicate mathematics learning difficulties and low mathematical achievement, respectively (Nelson & Powell, 2018). The 50th percentile was chosen to represent median performance and 75th and 90th to replicate the same distribution at the higher and lower ends. Model estimates between quantiles were compared using Analysis of Variance (ANOVA). The *quantreg* package in R (Koenker, 2020) was used to conduct these analyses.

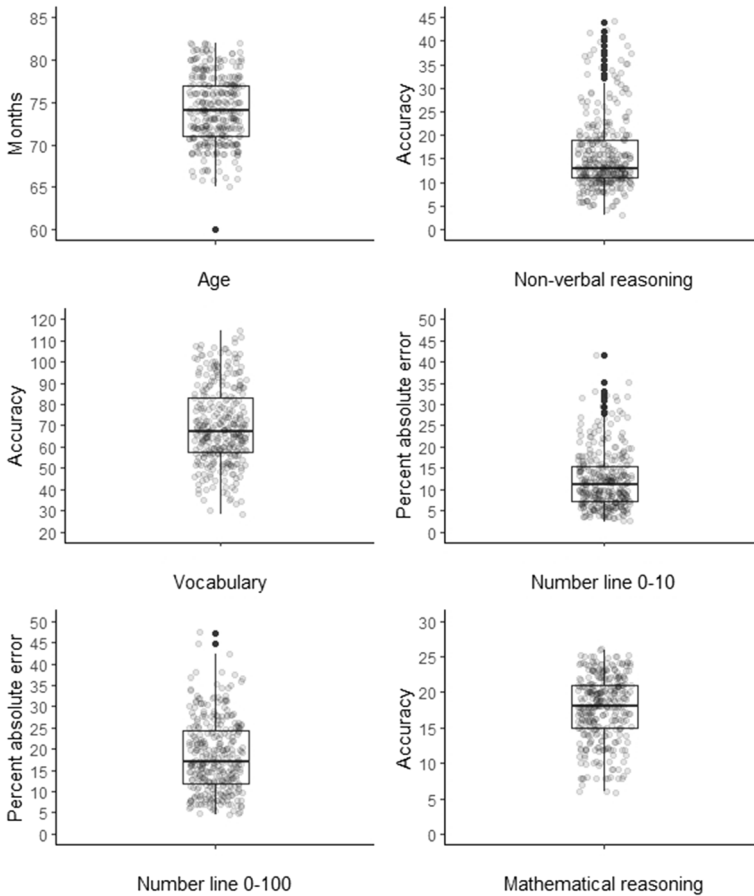
Descriptive statistics are presented in Table 1 with accompanying boxplots (Fig. 1). Pearson correlations between continuous variables included in the regression models are

Table 1 Descriptive Statistics

Variables	Mean	SD	Min	Max	Skewness	Kurtosis
Age	74.14	3.97	60	82	-0.05	2.53
Non-verbal reasoning	15.77	7.75	3	44	1.40	4.79
Vocabulary	70.02	18.13	28	115	0.31	2.48
Mathematical reasoning	17.81	4.54	6	26	-0.52	2.61
Number line 0–10	12.29	6.88	2.34	41.46	1.17	4.32
Number line 0–100	18.38	8.38	4.67	47.20	0.57	2.81

Age is presented in months. For number line estimation smaller values represent better performance. $N=324$

¹ Following standard practice, we refer to predictor when referring to the independent variables included in the linear and quantile regression analyses. However, these analyses are based on cross-sectional data, which allows us to report on concurrent associations between variables. By the use of the term "predictor" we are not implying causal relationships.



Boxplots represent the performance distribution of all continuous variables included in the analyses

Fig. 1 Boxplots for all Continuous Variables Included

presented in Table 2. As normative data with Singaporean samples is not available for the measures used, we report on raw scores. Before fitting the linear and quantile regressions, we calculated z-scores for all continuous variables. We first present the results of the linear regressions and then the quantile regressions.²

Number line's association with mathematical reasoning: linear regressions

To account for the effect of demographic characteristics and general cognitive skills on mathematical reasoning, Model 1 includes age, gender, non-verbal reasoning, and vocabulary. These were all significant predictors of mathematical reasoning. To study the

² We additionally conducted all analyses after controlling for outliers (i.e., setting all scores that were beyond 3 SD to values calculated at 3 SD from the mean – less than 1.5% of values for each variable). The results were consistent with the original analyses. Therefore, we only present the original analyses.

Table 2 Correlations between All Continuous Variables Included

Variables	1	2	3	4	5	6
1.Age	1	0.13	0.15*	0.22***	-0.03	-0.19**
2.Non-verbal reasoning		1	0.41***	0.42***	-0.18**	-0.31***
3.Vocabulary			1	0.46***	-0.15*	-0.35***
4.Mathematical reasoning				1	-0.29***	-0.51***
5.Number line 0–10					1	0.27***
6.Number line 0–100						1

For number line estimation smaller values represent better performance

$N=324$

* $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$

association of performance on the 0–10 number line task (Model 2) we added this independent variable to those present in Model 1. Performance on the 0–10 number line was significantly associated with mathematical reasoning and explained an additional 2.9% of unique variance ($F_{\text{change}}(1, 318) = 14.441, p < 0.001$).

To study the association of performance on the 0–100 number line task (Model 3) we added this independent variable to those present in Model 1. Performance on the 0–100 number line showed a significant association with mathematical reasoning and explained an additional 11.5% of unique variance ($F_{\text{change}}(1, 318) = 62.74, p < 0.001$). Full results for these models are presented in Table 3 and conditional scatterplots for number line estimation are presented in Figs. 2 and 3 for Models 2 and 3, respectively (see Supplementary Figures 1 and 2 for conditional scatterplots of all continuous predictors included in Models 2 and 3).

Model 3 shows better model fit than Model 2 ($\Delta AIC = 44$), and more variance in mathematical reasoning was explained by the 0–100 than the 0–10 number line (11.5% vs. 2.9%, respectively). We assessed for performance differences between both number line tasks through a Wilcoxon signed rank test as the normality assumption was not met. We found that children were significantly more accurate in the 0–10 than the 0–100 number line task ($p < 0.001$). We assessed if the association between performance on the number line task and mathematical reasoning differed according to the number range by using Zou's (2007) confidence interval method through the cocor package in R (Diedenhofen & Musch, 2015). We found that the correlation between mathematical reasoning and the 0–100 number line was significantly higher than the correlation between mathematical reasoning and the 0–10 number line ($r = -0.51$ and $r = -0.29$, 95% CI = [0.104, 0.336]). These findings seem to suggest a stronger association with mathematical reasoning of the 0–100 number line as compared to the 0–10 number line. We assessed if Model 2 was redundant in the presence of Model 3 by using the Cox Test and the Davidson–MacKinnon J test for comparing non-nested models through the lmttest package in R (Zeileis & Hothorn, 2002). We found that each model explained some variance in mathematical reasoning that was not accounted for by the other model ($p < 0.05$ for all comparisons). Therefore, we maintained Models 2 and 3 for the quantile regressions.

Number line's association with mathematical reasoning: quantile regressions

We retained Models 2 and 3 respectively to conduct quantile regressions (see Table 4). Performance on the 0–10 number line showed a significant association with mathematical

Table 3 Results of the Linear Regression Models

Predictors	Estimate	SE	<i>t</i>	<i>p</i>
Model 1				
Intercept	-0.107	0.066	-1.617	0.106
Gender	0.213	0.093	2.286	0.023
Age	0.128	0.047	2.694	0.007
Non-verbal reasoning	0.262	0.051	5.084	<0.001
Vocabulary	0.337	0.052	6.529	<0.001
<i>F</i>	$F(4, 319) = 34.83, p < 0.001$			
<i>Adj. R</i> ²	0.295			
AIC	813.1			
Model 2				
Intercept	-0.077	0.065	-1.178	0.240
Gender	0.153	0.093	1.653	0.099
Age	0.128	0.046	2.763	0.006
Non-verbal reasoning	0.235	0.051	4.618	<0.001
Vocabulary	0.322	0.051	6.338	<0.001
Number Line 0–10	-0.180	0.047	-3.800	<0.001
<i>F</i>	$F(5, 318) = 31.92, p < 0.001$			
<i>Adj. R</i> ²	0.324			
AIC	800.7			
Model 3				
Intercept	-0.160	0.061	-2.630	0.009
Gender	0.319	0.086	3.696	<0.001
Age	0.080	0.044	1.815	0.070
Non-verbal reasoning	0.190	0.048	3.957	<0.001
Vocabulary	0.242	0.049	4.968	<0.001
Number Line 0–100	-0.377	0.048	-7.921	<0.001
<i>F</i>	$F(5, 318) = 45.8, p < 0.001$			
<i>Adj. R</i> ²	0.410			
AIC	756.7			

All continuous variables included in the regressions were entered as *z* scores. For number line lower scores reflect better performance. Gender = reference category is boys

Fig. 2 Conditional Scatterplot for the effect of 0–10 Number Line on Mathematical Reasoning. The plot reflects the relationship between 0–10 number line and mathematical reasoning after accounting for the remaining predictors included in Model 2

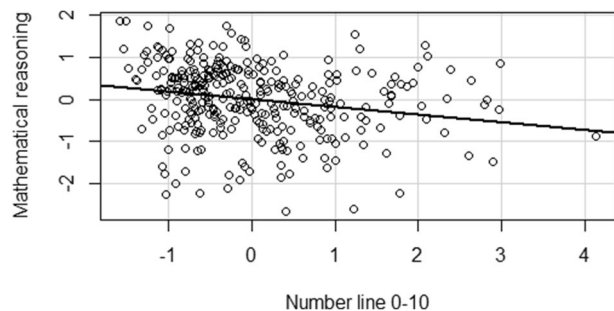
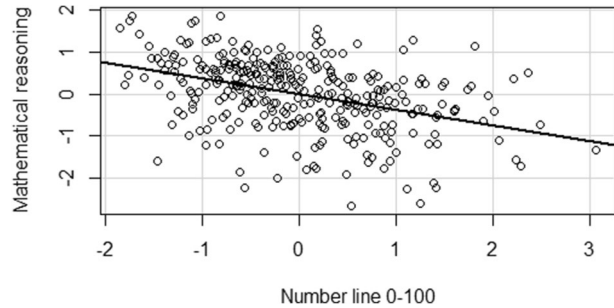


Fig. 3 Conditional Scatterplot for the Effect of 0–100 Number Line on Mathematical Reasoning. The plot reflects the relationship between 0–100 number line and mathematical reasoning after accounting for the remaining predictors included in Model 3



reasoning across all quantiles assessed. Slope comparisons between quantiles did not show a statistically significant difference ($F(4, 1616) = 0.507, p = 0.730$). Non-verbal reasoning and vocabulary were also significantly associated with mathematical reasoning across all quantiles. Slope comparisons showed statistically significant differences across quantiles for vocabulary only ($F(4, 1616) = 3.787, p = 0.005$). Post-hoc multiple comparisons (after Bonferroni adjustment) showed that the association of vocabulary and mathematical reasoning was significantly higher for the 25th than the 90th percentile ($\Delta\beta = 0.260, p = 0.028$). Figure 4 is the quantile coefficients plot for this model, which represents the slope coefficients for each predictor at each quantile examined accounting for all other predictors included in the model.

To isolate the effect of the number line from that of the other predictors included we also conducted quantile regression including only 0–10 number line as predictor. In this model 0–10 number line showed a significant association with ($p < 0.05$) mathematical reasoning for all quantiles included. Slope comparisons between quantiles showed a statistically significant difference between quantiles ($F(4, 1616) = 2.818, p = 0.024$). However, post-hoc multiple comparisons (after Bonferroni adjustment) were not statistically significant. Full results for this model appear in Supplementary Table 1.

Performance on the 0–100 number line showed a significant association with mathematical reasoning across all quantiles assessed. Slope comparisons between quantiles did not show a statistically significant difference ($F(4, 1616) = 1.309, p = 0.264$). Non-verbal reasoning and vocabulary were also significantly associated with mathematical reasoning across all quantiles. Slope comparisons did not show statistically significant differences across quantiles. Figure 5 is the quantile coefficients plot for this model, which represents the slope coefficients for each predictor at each quantile examined accounting for all other predictors included in the model.

As with the 0–10 number line, we also conducted quantile regression only including the 0–100 number line in the model. Performance on the 0–100 number line showed a significant association with mathematical reasoning across all quantiles assessed. Slope comparisons between quantiles did not show a statistically significant difference ($F(4, 1616) = 1.983, p = 0.095$). Full results for this model appear in Supplementary Table 1.

Discussion

The first aim of this study was to replicate the findings from previous research that show performance on the number line estimation task is significantly associated with mathematical reasoning. We replicated previous results finding that children that were more accurate

Table 4 Results from Quantile Regressions

	10 th percentile		25 th percentile		50 th percentile		75 th percentile		90 th percentile	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
For 0–10 number range										
Intercept	-1.136 ^{***}	0.175	-4986 ^{***}	0.105	0.060	0.061	0.447 ^{***}	0.064	0.786 ^{***}	0.115
Gender	0.098	0.191	0.012	0.130	0.045	0.111	0.238 ^{**}	0.081	0.329 ^{**}	0.118
Age	0.297 ^{**}	0.099	0.130	0.071	0.058	0.053	0.080	0.042	0.111 [*]	0.052
Non-verbal reasoning	0.408 ^{***}	0.109	0.234 ^{***}	0.053	0.272 ^{***}	0.054	0.183 ^{***}	0.030	0.158 ^{***}	0.047
Vocabulary	0.226 [*]	0.100	0.449 ^{***}	0.077	0.370 ^{***}	0.056	0.292 ^{***}	0.045	0.189 ^{***}	0.055
Number line 0–10	-0.183 ^{**}	0.068	-0.199 ^{**}	0.060	-0.215 ^{***}	0.057	-0.177 ^{***}	0.051	-0.122 [*]	0.056
AIC	1018.78		893.94		810.11		793.91		839.97	
For 0–100 number range										
Intercept	-1.26 ^{***}	0.132	-0.552 ^{***}	0.094	-0.035	0.064	0.377 ^{***}	0.065	0.687 ^{***}	0.097
Gender	0.349 [*]	0.139	0.256 [*]	0.127	0.292 ^{**}	0.098	0.342 ^{***}	0.070	0.313 ^{***}	0.088
Age	0.280 ^{***}	0.076	0.051	0.066	0.080	0.050	0.042	0.033	0.053	0.042
Non-verbal reasoning	0.277 ^{***}	0.083	0.238 ^{***}	0.067	0.206 ^{***}	0.046	0.140 ^{***}	0.029	0.103 ^{**}	0.034
Vocabulary	0.250 ^{**}	0.097	0.305 ^{***}	0.074	0.261 ^{***}	0.053	0.256 ^{***}	0.039	0.165 ^{***}	0.045
Number line 0–100	-0.371 ^{***}	0.055	-0.439 ^{***}	0.066	-0.421 ^{***}	0.056	-0.317 ^{***}	0.052	-0.309 ^{***}	0.054
AIC	988.15		855.73		760.22		748.48		799.39	

All continuous variables included in the regression were entered as z scores. Gender = reference category is boys. For number line estimation smaller values represent better performance

* $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$

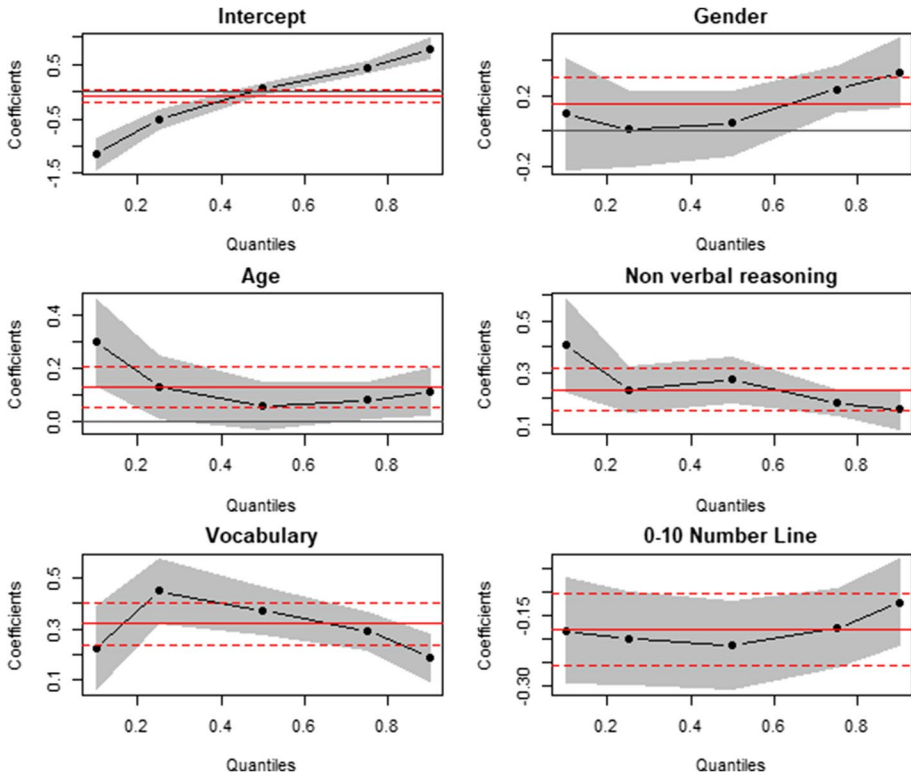


Fig. 4 Quantile Coefficients Plot with 0–10 Number Line as Predictor. The dots represent the slope coefficients (in y axis) for each quantile selected (in the x axis) and includes a line connecting them. The shaded bands represent the confidence intervals for the quantile regression coefficients at each quantile. The coefficient from the ordinary least squares regression is plotted in a full horizontal line with confidence intervals (discontinuous line). When zero is within the values in the y axis, it is represented by a full horizontal line

in the number line tasks tended to show better performance in mathematical reasoning for both number ranges. The regression estimate for the 0–100 number line ($\beta = -0.377$) in this study is similar to that of Zhu et al., (2017) with second graders ($\beta = -0.370$), but slightly higher than that found by Ouyang et al., (2021) with preschool children ($\beta = -0.275$). Performance on the 0–10 number line explained 2.9% of unique variance of mathematical reasoning while performance on the 0–100 number line explained 11.5% of unique variance. These findings suggest that, at a group level, our sample behaves similarly to those of previous studies.

The second aim was to determine if the relationship between number line estimation and mathematical reasoning holds across several points of the ability distribution. We found that for both number ranges, performance on the number line task showed a significant association with mathematical reasoning across all differing levels of performance. Even though the association tended to be higher for lower as compared to higher quantiles, we did not find statistically significant differences between quantiles. This is consistent with previous findings from Bartelet et al., (2014) who found that performance on a number comparison and an estimation task were associated with fact retrieval across the ability distribution.

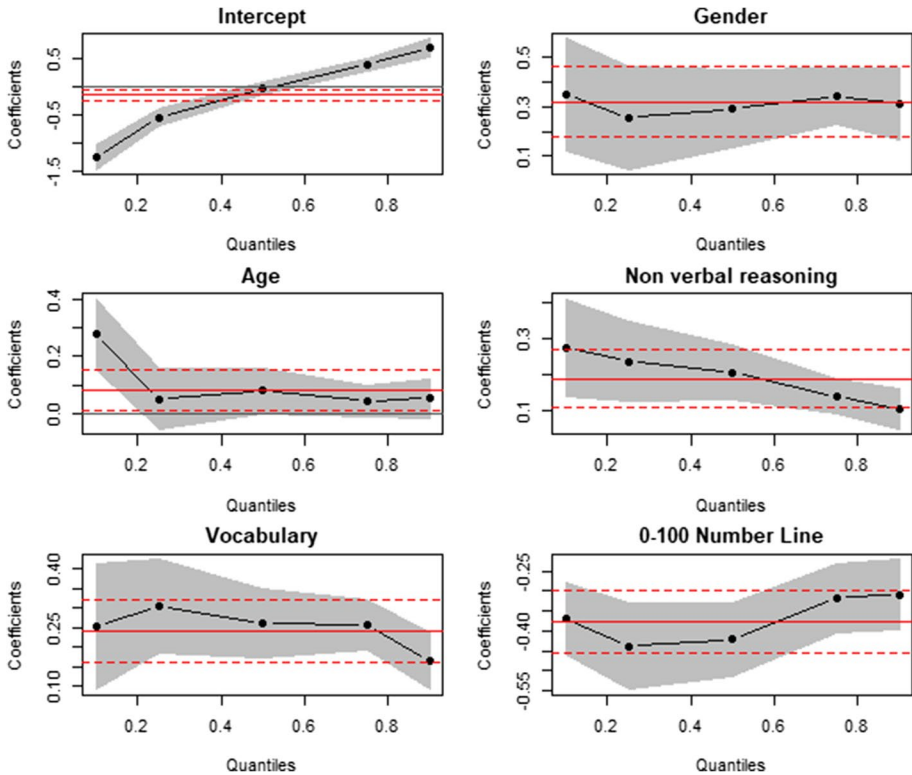


Fig. 5 Quantile Coefficients Plot for 0–100 Number Line as Predictor. The dots represent the slope coefficients (in y axis) for each quantile selected (in the x axis) and includes a line connecting them. The shaded bands represent the confidence intervals for the quantile regression coefficients at each quantile. The coefficient from the ordinary least squares regression is plotted in a full horizontal line with confidence intervals (discontinuous line). When zero is within the values in the y axis, it is represented by a full horizontal line

These findings add to the evidence that suggests that performance on the number line estimation task is strongly associated with concurrent mathematical reasoning and provides initial evidence that this is the case across all levels of ability. However, it is also possible that the homogeneity found in the relationship between number line estimation and mathematical reasoning is characteristic of younger children, and that when the demands of mathematical reasoning become more complex, further heterogeneity would arise. This would be consistent with some of Devlin et al., (2022) findings, as number relations was significantly predictive of mathematical achievement across all quantiles when assessed in pre-kindergarten but was only predictive at the 20th quantile in kindergarten and first grade. Future studies should expand to higher age ranges to explore if the relationship between performance on the number line estimation task and mathematical reasoning remains homogeneous in older children.

In the Introduction we proposed that differential relationships between performance on the number line task and mathematical reasoning might contribute to our understanding of what underlies performance on the number line estimation task and allow a closer approximation to understanding how this might support mathematical reasoning. However, as

significant differences between quantiles were not found, it is difficult to draw conclusions regarding what drives this relationship.

The number range of the number line task seems to play a role in the association between performance on the number line task and mathematical achievement, at least at the group level. Findings from the linear regression showed that at the group level, the association between number line estimation and mathematical reasoning was significantly higher for the 0–100 than the 0–10 number line task. These findings are consistent with the results of Ellis et al., (2021) who found that while both the 0–20 and 0–1000 number line tasks significantly predicted applied problems, the association was significantly stronger for the 0–1000 than the 0–20 number range. However, this is not consistent with Schneider et al., (2018) who found no effect of number range on the relationship between performance on the number line and mathematical achievement. Though the estimates for the 0–100 number line seem higher than those of the 0–10 number line ($0.140 < \Delta\beta < 0.240$) across all quantiles, we cannot know if these differences are of statistical significance, as to the best of our knowledge these cannot be formally compared.

Several reasons could explain the differences found between number ranges in the linear regression. For example, children could show more consistent performance on the number line estimation task with lower compared to higher number ranges. This would lead to overall less variability in performance on the number line estimation task with lower number ranges. This difference in variability across both number ranges is also reflected in the current sample (see Fig. 1), but it does not seem to be the only possible reason behind differences between number ranges. Relatedly, as suggested by Ellis et al., (2021), it is possible that when a number range becomes too easy, its relationship to mathematical ability might reduce. This has practical implications when using the number line estimation task in educational, research and clinical settings, as choosing a number range that is too easy for children might hinder the task's predictive capacity. Aligning the number range chosen to curriculum expectations is common practice (Link et al., 2014). However, our findings show that the number range that was aligned to curriculum expectations (i.e., 0–10), appears less predictive than the one that extended beyond them. This highlights the relevance of the number range chosen for the number line estimation task.

Alternatively, it is possible that as mentioned in the Introduction, performance on the number line task might index different skills depending on the number range being considered. If children are engaging with the 0–10 number line differently than with the 0–100 number line, then the differences between number ranges would reflect distinct underlying mathematical skills and their relationship with mathematical reasoning. Though this is a possibility, we would consider it more likely if variations in predictive strength between the number ranges were also accompanied with differences in the patterns found in the quantile regression. This was not the case in this sample, as the patterns of association between the number line estimation task and mathematical reasoning across the mathematical reasoning distribution were similar between both number ranges. These results suggest that the skills underlying performance on the number line task are significantly related to mathematical reasoning, though the number range included in the task could play an important role. Which role this might be remains unclear.

Findings from the linear and quantile regressions also highlighted the relevance of vocabulary and non-verbal reasoning for mathematical reasoning. These cognitive skills were significantly associated with mathematical reasoning across the ability distribution. Similar to performance on the number line task, the association of non-verbal reasoning and vocabulary tended to decrease as mathematical reasoning performance increased. However, significant differences between quantiles were only observed for vocabulary

between the 25th and 90th quantiles when performance on the 0–10 number line task was included in the model. Our findings regarding the role of vocabulary are not consistent with Fuchs et al., (2020) findings which showed that the association of language comprehension on mathematical reasoning was higher for higher levels of mathematical reasoning. However, our findings are consistent with those of Fuchs et al., (2020) who also found that non-verbal reasoning was a significant predictor of mathematical reasoning across all quantiles, with no significant differences between them. Overall, these findings highlight the relevance of considering both mathematical and non-mathematical skills when analyzing children's mathematical reasoning.

Implications

The fact that performance on the number line task predicted mathematical reasoning across its distribution highlights the potential of the number line estimation task as an assessment tool to provide relevant information for all children regardless of their level of ability. However, the number range chosen for the number line might impact this potential. This has practical implications when using the number line estimation task in educational, research and clinical settings, as choosing a number range that is too easy for children might hinder the task's predictive capacity. Aligning the number range chosen to curriculum expectations is common practice (Link et al., 2014). However, our findings suggest that the number range that was aligned to curriculum expectations (i.e., 0–10), appears less predictive than the one that extended beyond these expectations. This highlights the relevance of selecting a number range that is slightly challenging. Finally, our findings serve as a reminder that, at least when mathematical reasoning is concerned, it is crucial to consider the non-mathematical demands involved in mathematical reasoning. We remind the reader that the cross-sectional nature of this study does not allow to establish causal relations between performance on the number line estimation and mathematical reasoning. Thus, these implications should be read in the light of this limitation.

Limitations

One of the limitations of this study, as previously highlighted, is that conditional quantile regression considers the specified quantiles on the conditional distribution of the outcome variable. This means that the quantiles identified do not reflect the observed (unconditioned) performance distribution on the outcome variable, but its performance distribution dependent on (conditional on) the predictors included in the model (Wenz, 2019). Therefore, caution should be taken when practical implications are drawn from these analyses, as the percentiles identified based on quantile regression analyses might not completely overlap with those identified based on performance on mathematical reasoning alone. However, the findings from the models that only included performance on the number line task were consistent with those of the full models, strengthening the conclusions that can be drawn from the analyses presented. Relatedly, the identified quantiles are based on the performance of this sample. It is difficult to know to what extent the performance of this sample could be compared to that of other populations, as previous evidence suggests differences in the development of mathematical competence between Singaporean children and some international counterparts (Yao et al., 2017).

Additionally, we briefly present in the Introduction different theoretical positions that account for what performance on the number line task represents. However, this study

cannot provide evidence neither for nor against any of these accounts. Lack of consensus in the field in this regard, limits the possibility of fully interpreting the findings of this study in the context of what might underlie performance on the number line estimation task and how this relates to mathematical reasoning.

A final limitation of this study is its cross-sectional nature. Understanding how and why performance on the number line task relates to concurrent mathematical reasoning is a necessary first step. Practical educational implications, however, can only be drawn when we also understand how it relates to future mathematical reasoning and the acquisition of mathematical reasoning skills across all levels of ability.

Conclusion

The current study showed that performance on the 0–10 and 0–100 number line task is related to concurrent mathematical reasoning. This study also showed that even though, based on the curriculum, children included in this sample were only expected to know and manipulate numbers up to 10, performance on the 0–100 number line appears to be a stronger predictor of mathematical reasoning than performance in the 0–10 number line. The study did not find evidence to support that the relationship between performance on the number line tasks varies across differing levels of mathematical reasoning and highlights the potential of the number line estimation task for all children regardless of their level of ability. Finally, these findings add to previous evidence that suggests that the relationship between number line estimation and mathematical reasoning might differ depending on the number range being estimated.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1007/s10212-023-00708-2>.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions This manuscript was supported by an International Macquarie University Research Excellence Scholarship "iMQRES" Allocation No. 2020005 to the first author.

The original study was funded by the Singapore Ministry of Education (MOE) under the Education Research Funding Programme (OER 16/12RB) and administered by the National Institute of Education (NIE), Nanyang Technological University, Singapore.

Data availability The dataset for this study is stored at the Macquarie University Research Data Repository. Access to the data will be granted upon request. If you are interested in accessing the dataset please contact Rebecca Bull (r.bull@mq.edu.au). The code used for statistical analyses can be found in the project's OSF page (<https://osf.io/5prdy/>).

Declarations

Ethics approval The study received Ethics approval by the Institutional Review Board of Nanyang Technological University.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not

permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Australian Curriculum, Assessment and Reporting Authority [ACARA]. (2018). *Foundation to Year 10 curriculum (version 8.4)*. Available at <https://australiancurriculum.edu.au/f-10-curriculum>
- Ball, D., & Bass, H. (2003). Making mathematics reasonable in school. In J. Kilpatrick, G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 27–44). Reston, VA: National Council of Teachers of Mathematics.
- Bartelet, D., Vaessen, A., Blomert, L., & Ansari, D. (2014). What basic number processing measures in kindergarten explain unique variability in first-grade arithmetic proficiency? *Journal of Experimental Child Psychology*, *117*, 12–28. <https://doi.org/10.1016/j.jecp.2013.08.010>
- Bull, R., Lee, K., & Muñoz, D. (2021). Numerical magnitude understanding in kindergartners: A specific and sensitive predictor of later mathematical difficulties? *Journal of Educational Psychology*, *113*(5), 911–928. <https://doi.org/10.1037/edu0000640>
- Clarke, B., Strand Cary, M. G., Shanley, L., & Sutherland, M. (2020). Exploring the promise of a number line assessment to help identify students at-risk in mathematics. *Assessment for Effective Intervention*, *45*(2), 151–160. <https://doi.org/10.1177/1534508418791738>
- Cohen, D. J., & Blanc-Goldhammer, D. (2011). Numerical bias in bounded and unbounded number line tasks. *Psychonomic Bulletin and Review*, *18*, 331–338. <https://doi.org/10.3758/s13423-011-0059-z>
- Dackermann, T., Huber, S., Bahnmüller, J., Nuerk, H.-C., & Moeller, K. (2015). An integration of competing accounts on children's number line estimation. *Frontiers in Psychology*, *6*, 884. <https://doi.org/10.3389/fpsyg.2015.00884>
- Department for Education. (2014). *The national curriculum in England. Framework document*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/381344/Master_final_national_curriculum_28_Nov.pdf
- Department of Statistics Singapore. (2013). *Find data*. <https://www.singstat.gov.sg/find-data>
- Devlin, B. L., Jordan, N. C., & Klein, A. (2022). Predicting mathematics achievement from subdomains of early number competence: Differences by grade and achievement level. *Journal of Experimental Child Psychology*, *217*, 105354. <https://doi.org/10.1016/j.jecp.2021.105354>
- Dong, Y., Clements, D. H., Sarama, J., Dumas, D., Banse, H. W., & Day-Hess, C. (2020). Mathematics and executive function competencies in the context of interventions: A quantile regression analysis. *The Journal of Experimental Education*. <https://doi.org/10.1080/00220973.2020.1777070>
- Diedenhofen, B., & Musch, J. (2015). cocor: A Comprehensive Solution for the Statistical Comparison of Correlations. *PLoS ONE*, *10*(4), e0121945. <https://doi.org/10.1371/journal.pone.0121945>
- Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test-III*. American Guidance Service.
- Ebersbach, M., Luwel, K., & Verschaffel, L. (2015). The relationship between children's familiarity with numbers and their performance in bounded and unbounded number line estimations. *Mathematical Thinking and Learning*, *17*(2–3), 136–154. <https://doi.org/10.1080/10986065.2015.1016813>
- Ebersbach, M., Luwel, K., Frick, A., Onghena, P., & Verschaffel, L. (2008). The relationship between the shape of the mental number line and familiarity with numbers in 5- to 9-year-old children: Evidence for a segmented linear model. *Journal of Experimental Child Psychology*, *99*, 1–17. <https://doi.org/10.1016/j.jecp.2007.08.006>
- Education Endowment Foundation. (2017). *Improving mathematics in key stages 2 and 3. Guidance report*. Retrieved from https://d2tic4wvo1iusb.cloudfront.net/eef-guidance-reports/maths-ks-2-3/KS2_KS3_Maths_Guidance_2017.pdf
- Education Endowment Foundation. (2020). *Improving mathematics in the early years and key stage 1. Guidance report*. Retrieved from: https://d2tic4wvo1iusb.cloudfront.net/eef-guidance-reports/early-maths/EEF_Maths_EY_KS1_Guidance_Report.pdf
- Ellis, A., Susperreguy, M. I., Purpura, D. J., & Davis-Kean, P. E. (2021). Conceptual replication and extension of the relation between the number line estimation task and mathematical competence across seven studies. *Journal of Numerical Cognition*, *7*(3), 435–452. <https://doi.org/10.5964/jnc.7033>
- Fuchs, L. S., Newman-Gonchar, R., Schumacher, R., Dougherty, B., Bucka, N., Karp, K. S., Woodward, J., Clarke, B., Jordan, N. C., Gersten, R., Jayanthi, M., Keating, B., & Morgan, S. (2021). *Assisting students struggling with mathematics: Intervention in the elementary grades* (WWC 2021006). National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences,

- U.S. Department of Education. Retrieved from <https://ies.ed.gov/ncee/wwc/Docs/PracticeGuide/WWC2021006-Math-PG.pdf>
- Fuchs, L. S., Powell, S. R., Fall, A. M., Roberts, G., Cirino, P., Fuchs, D., & Gilbert, J. K. (2020). Do the processes engaged during mathematical word-problem solving differ along the distribution of word-problem competence? *Contemporary Educational Psychology*, *60*, 101811. <https://doi.org/10.1016/j.cedpsych.2019.101811>
- Hajovsky, D. B., Villeneuve, E. F., Schneider, W. J., & Caemmerer, J. M. (2020). An alternative approach to cognitive and achievement relations research: An introduction to quantile regression. *Journal of Pediatric Neuropsychology*, *6*, 83–95. <https://doi.org/10.1007/S40817-020-00086-3>
- Herbert, S., & Williams, G. (2023). Eliciting mathematical reasoning during early primary problem solving. *Mathematics Education Research Journal*, *35*, 77–103. <https://doi.org/10.1007/s13394-021-00376-9>
- Jeannotte, D., & Kieran, C. (2017). A conceptual model of mathematical reasoning for school mathematics. *Educational Studies in Mathematics*, *96*, 1–16. <https://doi.org/10.1007/s10649-017-9761-8>
- Koenker, R. (2020). *quantreg: Quantile regression*. R package version 5.75. <https://CRAN.R-project.org/package=quantreg>
- Li, H., Zhang, M., Wang, X., Ding, X., & Si, J. (2018). The central executive mediates the relationship between children's approximate number system acuity and arithmetic strategy utilization in computational estimation. *Frontiers in Psychology*, *9*, 943. <https://doi.org/10.3389/fpsyg.2018.00943>
- Lindmeier, A., Brunner, E., & Sumpter, L. (2018). Early mathematical reasoning – theoretical foundations and possible assessment In E. Bergqvist, M. Österholm, C. Granberg, & L. Sumpter (Eds.). *Proceedings of the 42nd conference of the international group for the psychology of mathematics education* (Vol. 3, pp. 315–322). Umeå, Sweden.
- Link, T., Huber, S., Nuerk, H. C., & Moeller, K. (2014). Unbounding the mental number line—new evidence on children's spatial representation of numbers. *Frontiers in Psychology*, *4*, 1021. <https://doi.org/10.3389/fpsyg.2013.01021>
- Lithner, J. (2008). A research framework for creative and imitative reasoning. *Educational Studies in Mathematics*, *67*(3), 255–276. <https://doi.org/10.1007/s10649-007-9104-2>
- Little, R. J. A. (1988). A test of Missing Completely at Random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*(404), 1198–1202. <https://doi.org/10.1080/01621459.1988.10478722>
- Ministry of Education. (2013). *Nurturing early learners: A curriculum for kindergartens in Singapore. Numeracy*. Retrieved from <https://www.nel.moe.edu.sg/qq1/slot/u143/Resources/Downloadable/pdf/nel-guide/nel-edu-guide-numeracy.pdf>
- Ministry of Education. (2021). *Mathematics syllabus. Primary one to six*. Author. Retrieved from: <https://www.moe.gov.sg/-/media/files/syllabus/2021-pri-mathematics-1-to-3.ashx>
- Moeller, K., Pixner, S., Kaufmann, L., & Nuerk, H.-C. (2009). Children's early mental number line: Logarithmic or decomposed linear? *Journal of Experimental Child Psychology*, *103*(4), 503–515. <https://doi.org/10.1016/j.jecp.2009.02.006>
- National Council for Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Author.
- Nelson, G., & Powell, S. R. (2018). A systematic review of longitudinal studies of mathematics difficulty. *Journal of Learning Disabilities*, *51*(6), 523–539. <https://doi.org/10.1177/0022219417714773>
- Nunes, T., Bryant, P., Evans, D., Bell, D., Gardner, S., Gardner, A., & Carraher, J. (2007). The contribution of logical reasoning to the learning of mathematics in primary school. *British Journal of Developmental Psychology*, *25*, 147–166. <https://doi.org/10.1348/026151006X153127>
- Nunes, T., Bryant, P., Barros, R., & Sylva, K. (2012). The relative importance of two different mathematical abilities to mathematical achievement. *British Journal of Educational Psychology*, *82*, 136–156. <https://doi.org/10.1111/j.2044-8279.2011.02033.x>
- OECD. (2018). *PISA 2021 Mathematics Framework (draft)*. <https://www.oecd.org/pisa/sitedocument/PISA-2021-mathematics-framework.pdf>
- Olsson, J., & Granberg, C. (2022). Teacher-student interaction supporting students' creative mathematical reasoning during problem solving using Scratch. *Mathematical Thinking and Learning*. Advance online publication. <https://doi.org/10.1080/10986065.2022.2105567>

- Ouyang, X., Yang, Y., Zhang, X., & Zhang, Q. (2021). Longitudinal relations between the approximate number system and symbolic number skills in preschool children. *Journal of Experimental Child Psychology*, 212, 105254. <https://doi.org/10.1016/j.jecp.2021.105254>
- Petscher, Y., & Logan, J. A. R. (2014). Quantile regression in the study of developmental sciences. *Child Development*, 85(3), 861–881. <https://doi.org/10.1111/cdev.12190>
- Podwysocki, C., Reeve, R. A., & Forte, J. D. (2019). The importance of ordinal information in interpreting number/letter line data. *Frontiers in Psychology*, 10, 692. <https://doi.org/10.3389/fpsyg.2019.00692>
- Purpura, D. J., & Logan, J. A. R. (2015). The nonlinear relations of the approximate number system and mathematical language to early mathematics development. *Developmental Psychology*, 51(12), 1717–1724. <https://doi.org/10.1037/dev0000055>
- Raven, J., Raven, J. C., & Court, J. H. (2003) *Manual for raven's progressive matrices and vocabulary scales*. Harcourt Assessment.
- R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Reinert, R. M., Hartmann, M., Huber, S., & Moeller, K. (2019). Unbounded number line estimation as a measure of numerical estimation. *PLoS ONE*, 14(3), e0213102. <https://doi.org/10.1371/journal.pone.0213102>
- Schneider, M., Merz, S., Stricker, J., De Smedt, B., Torbeyns, J., Verschaffel, L., & Luwel, K. (2018). Associations of number line estimation with mathematical competence: A meta-analysis. *Child Development*, 89(5), 1467–1484. <https://doi.org/10.1111/cdev.13068>
- Sella, F., Bertelletti, I., Lucangeli, D., & Zorzi, M. (2017). Preschool children use space, rather than counting, to infer the numerical magnitude of digits: Evidence for a spatial mapping principle. *Cognition*, 158, 56–67. <https://doi.org/10.1016/j.cognition.2016.10.010>
- Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, 75(2), 428–444. <https://doi.org/10.1111/j.1467-8624.2004.00684.x>
- Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology*, 62, 273–296. <https://doi.org/10.1016/j.cogpsych.2011.03.001>
- Sumpter, L. (2014). *Mathematical reasoning at pre-school level*. In: A mathematics education perspective on early mathematics learning between the poles of instruction and construction (POEM), Research symposium in Malmö, Sweden, Malmö University.
- Susperreguy, M. I., Davis-Kean, P. E., Duckworth, K., & Chen, M. (2018). Self-concept predicts academic achievement across levels of the achievement distribution: Domain specificity for math and reading. *Child Development*, 89(6), 2196–2214. <https://doi.org/10.1111/cdev.12924>
- Wechsler, D. (2005). *Wechsler Individual Achievement Test 2nd edition (WIAT II)*. The Psychological Corporation
- Wenz, S. E. (2019). What quantile regression does and doesn't do: A commentary on Petscher and Logan (2014). *Child Development*, 90(4), 1442–1452. <https://doi.org/10.1111/cdev.13141>
- Yackel, E., & Hanna, G. (2003). Reasoning and proof. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 227–236). Reston, VA: National Council of Teachers of Mathematics.
- Yao, S-Y., Muñoz, D., Bull, R., Lee, K., Khng, K. H., & Poon, K. (2017). Rasch modeling of the test of early mathematics ability—third edition with a sample of K1 children in Singapore. *Journal of Psychoeducational Assessment*, 35(6), 615–627. <https://doi.org/10.1177/0734282916651021>
- Yuan, L., Prather, R., Mix, K. S., & Smith, L. B. (2020). Number representations drive number-line estimates. *Child Development*, 91(4), e952–e967. <https://doi.org/10.1111/cdev.13333>
- Zeileis, A., & Hothorn, T. (2002). *Diagnostic checking in regression relationships*. *R News* 2(3), 7–10. <https://CRAN.R-project.org/doc/Rnews/>
- Zhu, M., Cai, D., & Leung, A. W. S. (2017). Number line estimation predicts mathematical skills: Difference in grades 2 and 4. *Frontiers in Psychology*, 8, 1576. <https://doi.org/10.3389/fpsyg.2017.01576>
- Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. *Psychological Methods*, 12(4), 399–413. <https://doi.org/10.1037/1082-989X.12.4.399>

Carola Ruiz

Current themes of research

Mathematics acquisition, number line estimation, mathematics learning difficulties, reading acquisition.

Relevant publications

- Simplicio, H., Gasteiger, H., Dorneles, B. V., Grimes, K. R., Haase, V. G., Ruiz, C., Liedtke, F. V., & Moeller, K. (2020). Cognitive Research and Mathematics Education—How Can Basic Research Reach the Classroom? *Frontiers in Psychology*, *11*(April), 1–5. <https://doi.org/10.3389/fpsyg.2020.00773> .
- Balbi, A., von Hagen, A., Cuadro, A. & Ruiz, C. (2019). Precursores de la Competencia Lectora Inicial en escolares Hispanoparlantes de nivel socioeconómico vulnerable [Precursors of Early Reading Competence in Spanish-Speaking Children from Vulnerable Backgrounds], *Psykhē*. *29*(1), 1-15. <https://doi.org/10.7764/psykhe.29.1.1403>.
- Ruiz, C. & Balbi, A. (2018) The effects of teaching mental calculation in the development of mathematical abilities. *The Journal of Educational Research*, *112* (3), 315-326. <https://doi.org/10.1080/00220671.2018.1519689> .
- Singer, V., Ruiz, C. & Cuadro, A. (2018) Las Habilidades lingüísticas y el sistema numérico aproximado en la eficacia en el cálculo aritmético. [The influence of linguistic abilities and the approximate number system in the efficacy of arithmetic calculation] *Bordón. Revista de Pedagogía*, *70* (3) 185-197. <https://doi.org/10.13042/Bordon.2018.62928>.
- Balbi, A., von Hagen, A., Cuadro, A. & Ruiz, C. (2018) Revisión sistemática sobre intervenciones en alfabetización temprana: implicancias para intervenir en español. [A systematic review on early literacy interventions: implications for interventions in Spanish] *Revista Latinoamericana de Psicología*, *50* (1), 31-48. <https://doi.org/10.14349/rlp.2018.v50.n1.4> .

Saskia Kohnen

Current themes of research

Reading acquisition, spelling acquisition, dyslexia, mathematics acquisition.

Relevant publications

- Colenbrander, D., Kohnen, S., Beyersmann, E., Robidoux, S., Wegener, S., Arrow, T., ... & Castles, A. (2022). Teaching Children to Read Irregular Words: A Comparison of Three Instructional Methods. *Scientific Studies of Reading*, 1-20. <https://doi.org/10.1080/10888438.2022.2077653>.
- von Hagen, A., Kohnen, S., & Stadie, N. (2021). Foreign language attainment of children/adolescents with poor literacy skills: A systematic review and meta-analysis. *Educational Psychology Review*, *33*(2), 459-488. <https://doi.org/10.1007/s10648-020-09566-6> .
- Francis, D., Hudson, J. L., Kohnen, S., Mobach, L., & McArthur, G. M. (2021). The effect of an integrated reading and anxiety intervention for poor readers with anxiety. *PeerJ*, *9*, e10987. <https://doi.org/10.7717/peerj.10987> .
- Kohnen, S., Banales, E., & McArthur, G. (2021). Videoconferencing interventions for children with reading and spelling difficulties: a pilot study. *Telemedicine and e-Health*, *27*(5), 537-543. <https://doi.org/10.1089/tmj.2020.0061> .
- Larsen, L., Schaubert, S. K., Kohnen, S., Nickels, L., & McArthur, G. (2020). Children's knowledge of single-and multiple-letter grapheme-phoneme correspondences: An exploratory study. *Early Childhood Research Quarterly*, *51*, 379-391. <https://doi.org/10.1016/j.ecresq.2019.12.001>.
- Wheldall, K., Wheldall, R., Madelaine, A., Reynolds, M., Arakelian, S., & Kohnen, S. (2019). 'Just Teach Our Kids to Read': Efficacy of Intensive Reading Interventions for Both Younger and Older Low-Progress Readers in Schools Serving Mainly Remote Indigenous Communities. In *Literacy Education and Indigenous Australians* (pp. 221-246). Springer, Singapore. https://doi.org/10.1007/978-981-13-8629-9_13.
- Larsen, L., Kohnen, S., McArthur, G., & Nickels, L. (2018). An investigation of grapheme parsing and grapheme-phoneme knowledge in two children with dyslexia. *Reading and writing*, *31*(4), 991-1015. <https://doi.org/10.1007/s11145-018-9823-z>.

- Brawn, G., Kohnen, S., Tassabehji, M., & Porter, M. (2018). Functional basic reading skills in Williams syndrome. *Developmental Neuropsychology*, 43(5), 454-477. <https://doi.org/10.1080/87565641.2018.1455838>.
- McArthur, G., Sheehan, Y., Badcock, N. A., Francis, D. A., Wang, H. C., Kohnen, S., ... & Castles, A. (2018). Phonics training for English-speaking poor readers. *Cochrane Database of Systematic Reviews*, (11). <https://doi.org/10.1002/14651858.CD009115.pub3>.
- Kohnen, S., Nickels, L., Geigis, L., Coltheart, M., McArthur, G., & Castles, A. (2018). Variations within a subtype: Developmental surface dyslexias in English. *Cortex*, 106, 151-163. <https://doi.org/10.1016/j.cortex.2018.04.008>.
- Marinus, E., Kezilas, Y., Kohnen, S., Robidoux, S., & Castles, A. (2018). Who are the noisiest neighbors in the hood? Using error analyses to study the acquisition of letter-position processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(9), 1384. <https://doi.org/10.1037/xlm0000524>.

Rebeca Bull

Current themes of research

Mathematics acquisition, mathematics learning difficulties, early childhood education, mathematical cognition, executive functioning.

Relevant publications

- Muñoz, D., Bull, R., & Lee, K. (2022). Maternal education and siblings: Agents of cognitive development in kindergarten. *Developmental Science*, 25(4), e13218. <https://doi.org/10.1111/desc.13218>.
- Muñoz, D., Bull, R., Cheung, P., & Orrantia, J. (2022). Non-symbolic ratio reasoning in kindergartners: Underlying unidimensional heuristics and relations with math abilities. *Frontiers in Psychology*, 13, 800977. <https://doi.org/10.3389/fpsyg.2022.800977>.
- Muñoz, D., Lee, K., Bull, R., Khng, K. H., Cheam, F., & Rahim, R. A. (2022). Working memory and numeracy training for children with math learning difficulties: Evidence from a large-scale implementation in the classroom. *Journal of Educational Psychology*. Advance online publication. <https://doi.org/10.1037/edu0000732>.
- Wong, B., Bull, R., Ansari, D., Watson, D. M., & Liem, G. A. D. (2022). Order processing of number symbols is influenced by direction, but not format. *Quarterly Journal of Experimental Psychology*, 75(1), 98-117. <https://doi.org/10.1177/17470218211026800>.
- Johnston, K., & Bull, R. (2022). Understanding educator attitudes towards and perceptions of mathematics in early childhood. *Journal of Early Childhood Research*, 20(3), 341-356. <https://doi.org/10.1177/1476718X211052785>.
- Ng, E. L., Bull, R., Bautista, A., & Poon, K. (2021). A bifactor model of the Classroom Assessment Scoring System in preschool and early intervention classrooms in Singapore. *International Journal of Early Childhood*, 53(2), 197-218. <https://doi.org/10.1007/s13158-021-00292-w>.
- Ng, E., Bull, R., & Khng, K. H. (2021, September). Accounting for the SES-math achievement gap at school entry: Unique mediation paths via executive functioning and behavioral self-regulation. *Frontiers in Education*, 6, 703112. <https://doi.org/10.3389/educ.2021.703112>.
- Kwok, F. Y., Bull, R., & Muñoz, D. (2021). Cross-and within-domain associations of early reading and mathematical skills: Changes across the preschool years. *Frontiers in Psychology*, 12, 710470. <https://doi.org/10.3389/fpsyg.2021.710470>.
- Bull, R., Lee, K., & Muñoz, D. (2021). Numerical magnitude understanding in kindergartners: A specific and sensitive predictor of later mathematical difficulties? *Journal of Educational Psychology*, 113(5), 911-928. <https://doi.org/10.1037/edu0000640>.
- Muñoz, D., Bull, R., & Lee, K. (2021). Socioeconomic status, home mathematics environment and math achievement in kindergarten: A mediation analysis. *Developmental Science*, 24(6), e13135. <https://doi.org/10.1111/desc.13135>.
- Bautista, A., Bull, R., Ng, E. L., & Lee, K. (2021). "That's just impossible in my kindergarten." Advocating for 'glocal' early childhood curriculum frameworks. *Policy Futures in Education*, 19(2), 155-174. <https://doi.org/10.1177/1478210320956500>.
- Bautista, A., Moreno-Núñez, A., Vijayakumar, P., Quek, E., & Bull, R. (2020). Gross motor teaching in preschool education (Enseñanza de la motricidad gruesa en educación infantil): where, what and how

- do Singapore educators teach?(¿ dónde, qué y cómo enseñan las maestras en Singapur?). *Infancia y aprendizaje*, 43(2), 443-482. <https://doi.org/10.1080/02103702.2019.1653057>.
- Bautista, A., Habib, M., Ong, R., Eng, A., & Bull, R. (2019). Multiple Representations in Preschool Numeracy: Teaching a Lesson on More-or-Less. *Asia-Pacific Journal of Research in Early Childhood Education*, 13(2), 95-122. <https://doi.org/10.17206/apjrece.2019.13.2.95>.
- Bautista, A., Habib, M., Eng, A., & Bull, R. (2019). Purposeful play during learning centre time: From curriculum to practice. *Journal of Curriculum Studies*, 51(5), 715-736. <https://doi.org/10.1080/00220272.2019.1611928>.
- Bull, R., & Bautista, A. (2018). A careful balancing act: Evolving and harmonizing a hybrid system of ECEC in Singapore. In S. L. Kagan (Ed.), *The Early Advantage: Early Childhood Systems that Lead by Example* (pp. 155-181). New York: Teachers College Press.
- Ng, S. C., & Bull, R. (2018). Facilitating social emotional learning in kindergarten classrooms: Situational factors and teachers' strategies. *International Journal of Early Childhood*, 50(3), 335-352. <https://doi.org/10.1007/s13158-018-0225-9>.
- Lee, K., Ng, S. F., & Bull, R. (2018). Learning and solving algebra word problems: The roles of relational skills, arithmetic, and executive functioning. *Developmental psychology*, 54(9), 1758. <https://doi.org/10.1037/dev0000561>.
- Bull, R., Marschark, M., Nordmann, E., Sapere, P., & Skene, W. A. (2018). The approximate number system and domain-general abilities as predictors of math ability in children with normal hearing and hearing loss. *British Journal of Developmental Psychology*, 36(2), 236-254. <https://doi.org/10.1111/bjdp.12204>.