Journal of Numerical Cognition

PsychOpen GOLD

Empirical Research

Check for updates

Number Line Estimation Patterns and Their Relationship With Mathematical Performance

Carola Ruiz [1], Saskia Kohnen [2], Rebecca Bull [1]

[1] *School of Education, Macquarie University, Sydney, Australia.* **[2]** *School of Psychological Sciences, Macquarie University, Sydney, Australia.*

Journal of Numerical Cognition, 2023, Vol. 9(2), 285–301, https://doi.org/10.5964/jnc.10557

Received: 2022-10-30 • **Accepted:** 2023-04-14 • **Published (VoR):** 2023-07-31

Handling Editor: Lieven Verschaffel, KU Leuven, Leuven, Belgium

Corresponding Author: Carola Ruiz, 2 First Walk, Macquarie University, NSW2109, Australia. E-mail: carola.ruizhornblas@hdr.mq.edu.au

Supplementary Materials: Materials [see Index of Supplementary Materials]

Abstract

There is ongoing debate regarding what performance on the number line estimation task represents and its role in mathematics learning. The patterns followed by children's estimates on the number line task could provide insight into this. This study investigates children's estimation patterns on the number line task and assesses whether mathematics achievement is associated with these estimation patterns. Singaporean children (n = 324, Age M = 6.2 years, Age SD = 0.3 years) in their second year of kindergarten were assessed on the number line task (0-100) and their mathematical performance (Numerical Operations and Mathematical Reasoning subtests from WIAT II). The results show that most children's number line estimation patterns can be explained by at least one mathematical model (i.e., linear, logarithmic, unbounded power model, one-cycle power model, two-cycle power model). But the findings also highlight the high percentage of participants for which more than one model shows similar support. Children's mathematical achievement differed based on the models that best explained children's estimation patterns. Children whose estimation patterns corresponded to a more advanced model tended to show higher mathematical achievement. Limitations of drawing conclusions regarding what performance on the number line task represents based on models that best explain the estimation patterns are discussed.

Keywords

estimation patterns, number line, representational shift, proportional judgment, mathematical achievement

Performance on the number line estimation task has the potential to provide crucial information regarding mathematics learning. However, disagreements on what is reflected by performance on the number line estimation task, and what role it might play in mathematics learning remain. Previous findings highlight the relationship between accuracy on the number line task and other mathematical skills (see Schneider et al., 2018 for a meta-analysis). For example, intervention studies have shown that accuracy on the number line estimation task is trainable and that it can impact other mathematical skills (Dackermann et al., 2016; Elofsson et al., 2016; Friso-van den Bos et al., 2018; Kucian et al., 2011; Link et al., 2013; Maertens et al., 2016; Ramani & Siegler, 2011; Siegler & Ramani, 2009; Whyte & Bull, 2008). Furthermore, longitudinal studies have highlighted accuracy on the number line task as a predictor of later low mathematical achievement (Bull et al., 2021). However, less research has focused on the patterns of children's estimates on the number line task and how these relate to performance on other mathematical skills, while considering different theoretical interpretations. The aim of this study is to examine children's estimation patterns on the number line task and assess whether mathematics achievement is associated with these estimation patterns. The study aims are addressed from two theoretical accounts, which are first considered independently and then compared to each other. This dual approach is taken to better understand if the approach taken impacts the interpretation of what performance on the number line represents and how it relates to mathematical achievement.

The number line task, in its traditional version, consists of a horizontal line with an upper and lower limit. These limits are labelled and indicate the range of the number line (e.g., 0-10, 0-100). The task consists of placing Arabic numbers on the horizontal line based on the number's value. Accuracy is represented by error measures based on the distance between the estimate and the target number, which are then corrected for the range of the number line (i.e., percent absolute error, PAE). An alternative way to index performance involves mapping the estimates (on the y axis) against the target numbers (on the x axis) to then assess the fit of different mathematical functions to the patterns formed. These patterns are usually referred to as estimation patterns, and what is being assessed is which mathematical function most closely resembles (i.e., explains) the estimates. Estimation patterns can be assessed at a group-level, which involves combining the estimates' placements across participants (but see Bouwmeester & Verkoeijen, 2012 for a critique of this approach) or at an individual level. In this study we only consider estimation patterns at an individual level.

The link between estimation patterns and several mathematical functions have been studied (e.g., linear, logarithmic, unbounded, one and two-cycle power models). There is evidence to suggest that the patterns followed by children's estimates may shed some light on what performance on the number line task reflects, as well as the way in which children engage with the task (see Dackermann et al., 2015; Yuan et al., 2020; for a complete overview). In this study, we will focus on two of the theoretical accounts that have been extensively discussed in the literature, logarithmic to linear representational shift (representational shift from here onwards) and proportional judgment.

Representational Shift Account

This theoretical account proposes that the mental representation of magnitude shifts from logarithmic to linear (Kim & Opfer, 2017) across development, and that performance on the number line estimation task acts as a direct reflection of the person's mental representation of magnitude (Siegler & Booth, 2005; but see Thompson et al., 2022 for a more nuanced view of this assumption). This shift from logarithmic to linear is dependent on age (Kim & Opfer, 2017; Siegler & Booth, 2004) and the number range (Opfer & Siegler, 2007), and has usually been referred to as the 'log to linear shift' (Opfer et al., 2016). This means that we would expect estimation patterns to better resemble a logarithmic function for younger children (or for older children when estimating on less known number ranges). In contrast, the estimation patterns of older children and adults (or younger children working with very-well known number ranges) would be expected to fit a linear function. When the mental representation of magnitude resembles a logarithmic function, the distances between consecutive numbers are perceived as bigger for lower than for higher numbers (i.e., the distance between 5 and 6 is perceived as bigger than the distance between 55 and 56). In contrast, when the mental representation of numbers resembles a linear function, the distance between consecutive numbers is maintained constant throughout the number range. Thus, estimation patterns on the number line task shifting from logarithmic to linear would be driven by a shift in the person's mental representation of magnitude. However, the assumption that performance on this task directly reflects the mental representation of magnitude has received some criticism (Bouwmeester & Verkoeijen, 2012; Huber et al., 2014).

Numerous studies have provided evidence for estimation patterns evolving from logarithmic to linear. Differences between studies' methodologies and findings make it difficult to determine when this shift is expected to occur within any given number range. For example, a logarithmic model best fit the estimation patterns of over 80% of four- to six-year-old preschoolers on a 0-1000 number line (Yuan et al., 2020), and of six-year-old kindergarteners on a 0-100 number line (Siegler & Booth, 2004). But, in Praet and Desoete's (2014) study only half of kindergarteners' estimation patterns on the 0-100 number line were best fit by a logarithmic model with over 40% of estimates that could not be fitted by a logarithmic nor a linear model. Studies with first graders found that a logarithmic model best fit sightly over 60% of estimation patterns on a 0-100 number line, while a linear model best fit estimation patterns of approximately 30% of children (Praet & Desoete, 2014; Sasanguie et al., 2016; Siegler & Booth, 2004). In second grade, the linear model became the best fit for between 55% (Praet & Desoete, 2014; Siegler & Booth, 2004) and 70% (Sasanguie et al., 2016) of estimation patterns on the 0-100 number line, though the logarithmic model still best fit between 30% (Praet & Desoete, 2014; Sasanguie et al., 2016) and 45% (Siegler & Booth, 2004) of estimation patterns. These findings highlight individual differences regarding when children produce linear estimates within a given number range. Though there is some evidence that this could occur during the preschool years, it seems more likely that the shift from logarithmic to linear on the 0-100 number line is ongoing for at least the first two years of primary school.

Proportional Judgment Account

The proportional judgment account proposes that performance on the number line estimation task requires the estimation of the proportion between two numbers (the target number and the number range). This is in line with psychophysical models of relative quantity judgments within a proportion judgment framework (see Hollands & Dyre, 2000 for an explanation of this framework). As performing the number line task involves considering more than one number simultaneously, within this theoretical account estimation patterns cannot be considered as a direct reflection of the mental representation of magnitude (Barth & Paladino, 2011). According to this theoretical account, performance on the number line task is impacted by two factors (i.e., sources of change): the use of benchmarks and estimation bias. This account suggests that the number line task is completed by using benchmarks (Peeters et al., 2017), which can be provided (i.e., limits of the number line) or are internal (i.e., estimated by the person). At the lower end of proficiency, children might only be able to use the start point of the number line. In this case, the number line task is solved as a single magnitude judgment task, and the estimation pattern would resemble an unbounded power function, adapted from Stevens' power law ($y = a\left(X^b\right)$, Stevens, 1975). More advanced performance would follow for those able to use both start and end points. In this case, the estimation pattern is expected to resemble a one-cycle power model characterized by an S (or inverse S) shaped pattern. Additionally, some people can use the limits of the number line while also inferring a midpoint. The estimation pattern in this case is expected to resemble a two-cycle power model characterized by two repetitions of the S shaped pattern (i.e., one before and one after the inferred midpoint; see Slusser et al., 2013 for an explanation of these models). The second source of change is estimation bias. The estimation of any magnitude is accompanied by a certain level of bias, which depends on the type of magnitude being estimated (Hollands & Dyre, 2000). However, as mentioned before, the estimation in a bounded number line task entails the estimation of two magnitudes. This would mean that the estimation bias would encompass the bias associated with each of the numbers (Slusser et al., 2013). The level of estimation bias is reflected by β, which is a free parameter included in all three models. Most commonly β is lower than 1, this represents a pattern of over and then under estimation, though the opposite pattern (i.e., under and then over estimation) is possible, and is reflected by β values above 1. Values of β = 1 correspond to no estimation bias (i.e., perfectly linear estimates; Hollands & Dyre, 2000). As children develop, the level of bias decreases (Slusser et al., 2013; Zax et al., 2019), making the estimates more accurate (β values move towards 1). From this perspective, the evolution of estimation patterns reflects a decrease in estimation bias together with a shift in the number of benchmarks used to solve the task.

As with the studies discussed in the previous section, determining the age at which the transitions between number of benchmarks used are expected to occur for a particular number range is challenging. For example, three- and four-year-olds' estimation patterns on a 0-10 number line were best explained by an unbounded power model (39.4%), followed by a one-cycle power model (18.3%), and finally a two-cycle power model (11%). Approximately one third of estimation patterns could not be fitted by any of these three models (Liang et al., 2021). Six-to-eight-year olds' estimation patterns on a 0-100 number line were best fit by a one-cycle power model (61.2%), followed by the unbounded power model and the two-cycle power model (19.4% each, Xing et al., 2021). However, the authors found an age effect, where younger children's estimation patterns were more likely to be best fit by an unbounded power model, and older children's estimation patterns more frequently being best fit by the two-cycle power model. Similarly, Sasanguie et al. (2016) found that most first (53%), second (between 56% and 64% depending on the sample), and third graders' estimation patterns (72%) on a 0-100 number line were best fit by either a one or two-cycle power model (without distinguishing between them), with the remaining children's estimation patterns being best fit by the unbounded power model. These findings highlight the variability found within any given age group and number range.

However, some patterns seem to emerge with an increasing percentage of children using more advanced strategies, and therefore more benchmarks to solve the number line task, as they get older.

Challenges Involved in the Study of Estimation Patterns

Previous studies have found evidence for and against models arising from both theoretical accounts. These differing findings may occur due to methodological differences and existing challenges for the study of estimation patterns. One of the main challenges is the lack of an absolute reference for model selection. This means that a given mathematical function is the one that best explains the estimation patterns only in comparison with the others being fitted to the same estimation patterns. Another challenge arises when considering the number and types of models that are compared. So far, we have only discussed comparisons between models arising from a single theoretical account. However, findings get more complex when studies compare models that arise from different theoretical accounts. For example, in Slusser et al. (2013), five- and six-year-olds' estimation patterns on a 0-20 number line were distributed as follows: unbounded power model and linear model (30% each), one-cycle power model (25%) and two cycle power model (15%). In Xu et al. (2013), six-year-old preschoolers' estimation patterns on a 0-100 number line showed the following distribution: two-cycle power model (22%), linear model (13.6%), and one-cycle power model (5.1%). No estimation patterns were best fit by a logarithmic model, and the remaining estimation patterns were best fit by a two-linear model. Link et al. (2014) found that a linear model was most frequently the best fitting model for first graders (64%) on a 0-10 number line, followed by an unbounded power model (21%), a one-cycle power model (9%) and finally a two-cycle model (6%). In this same study, second graders' estimation patterns on a 0-20 number line showed a similar distribution: linear (51%), unbounded power model (23%), one-cycle power model (12%) and two-cycle power model (9%). Finally, Sasanguie et al. (2016) found that models arising from the proportional judgment account best explained the estimation patterns of 53% of first graders, between 59 and 75% of second graders, and of 69% of third graders, while linear and logarithmic models best explained the estimation patterns of 40% of first graders, between 21 and 35% of second graders, and 20% of third graders. It seems clear that the variability found in children's estimation patterns is confounded by the number of models being compared.

A related challenge, which is rarely made explicit, is the slight differences that are sometimes found in model fit between models. This entails that a model is often chosen over another one, without strong evidence for its preference. The fact that this is rarely reported (at least related to individual children's estimation patterns) makes it harder to know how prevalent it might be. These limitations complicate our understanding of what estimation patterns on the number line task represent and how they relate to other mathematical abilities.

Estimation Patterns and Their Relationship With Mathematical Achievement

Accuracy on the number line estimation task has been related to concurrent and future mathematical achievement (Schneider et al., 2018). Few studies have explored the relationship between children's estimation patterns and their mathematical achievement. For example, Friso-van den Bos et al. (2015) found that throughout first grade, children whose estimation patterns were best fit by a linear or a one-cycle power model tended to outperform those who were best fit by logarithmic and non-cyclic power models on measures of mathematics achievement. Xing et al. (2021) found that six- to eight-year-olds whose estimation patterns on a 0-100 number line were best fit by a two-cycle power model outperformed those best fit by a one-cycle and an unbounded power model on a standardized mathematics test. These findings help extend our understanding regarding how estimation patterns relate to other mathematical skills, but more research that considers multiple theoretical accounts is needed.

The Present Study

We approach the aims of this study based on the representational shift and proportional judgment accounts, first considered separately and then concurrently. The first aim is to determine which mathematical functions (i.e., linear, logarithmic, unbounded power, one-cycle power, and two-cycle power) best explain children's estimation patterns on the 0-100 number line task. The second aim is to determine if and how children's mathematical skills are associated with the mathematical function that best explains their estimation patterns.

Only a handful of previous studies have considered both theoretical accounts when analysing individual estimation patterns (Barth & Paladino, 2011; Friso-van den Bos et al., 2015; Link et al., 2014; Sasanguie et al., 2016; Slusser et al., 2013; Xu et al., 2013). Of these, only Sasanguie et al. (2016) performed their analysis for the representational shift and proportional judgment account separately, and then compared both accounts to each other. However, in this study, when considering the proportional judgment account, one and two-cycle power models were combined under the same group, which does not allow to differentiate between these two. Furthermore, when comparing both theoretical accounts, Sasanguie et al. reported on the percentage of participants for which each theoretical account best explained their estimation patterns, but not on the percentage of estimation patterns best explained by each model. For example, 33 percent of children's estimation patterns were best explained by one of the models from the representational shift account when compared to the models from the proportional judgment account, but how much of that 33 percent corresponded to the linear model and how much to the logarithmic model was not stated.

Even fewer studies have looked at the relationship between children's individual estimation patterns and their mathematical skills (Friso-van den Bos et al., 2015; Xing et al., 2021). Only Friso-van den Bos et al. (2015) included both theoretical accounts. However, as was the case with most studies mentioned earlier, the Friso-van den Bos study compared all models to each other, without including an analysis based on each theoretical account individually. Therefore, an analysis of children's individual estimation patterns and their relationship to mathematics achievement which considers all models arising from each of these theoretical accounts separately, as well as a comparison between theoretical accounts is needed. This approach might provide a better understanding of what underlies performance on the number line estimation task.

Materials and Method

Participants

Singaporean children (n = 347) in their second year of Kindergarten (year before starting primary school) were recruited. This sample represents 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). Data was found to be likely missing completely at random, $\chi^2(24) = 29.445$, $p = .204$, (Little, 1988). Listwise deletion was used, and 23 observations were deleted from the dataset. Three hundred and twenty-four participants composed the final sample (162 boys, $M_{age} = 6.2$ years, SD_{age} = 0.3 years). Participants' ethnicity as reported by their parents was distributed as follows: 63.0% Chinese, 21.0% Malay, 9.3% Indian and 5.2% from other ethnicities. Ethnicity for five children was not reported. A comparison with census data indicates an under-representation of Chinese children and an over-representation of Malay children (census = 74.2% and 13.3%, respectively; Department of Statistics Singapore, 2013). Parents' highest educational level achieved was reported for 70% of the sample. Parents without a high school qualification were underrepresented in the sample (8.3% vs. 31.2% from census data), while other qualifications mostly aligned with census data − 17.3% had high school qualifications, 22.2% had post-secondary qualifications, and 22.6% had a university undergraduate or postgraduate qualification.

Procedure

Parents provided written consent, and ethical approval by the Institutional Review Board at Nanyang Technological University was obtained. Testing was conducted in two to three sessions, held on different days. All children were tested in their preschools, in the language of instruction (English), and given a small gift for participating.

Measures

Number Line Estimation

Children solved a computerised bounded number line task (0-100). Children were presented with a horizontal line with zero at the left and 100 at the right and were shown a number above the middle of the line which they had to place on the line. Children initially completed one practice trial (fifty) to confirm that they had understood the task and were then assessed with 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). PAE was used to index accuracy, calculated as [(number estimated - target number)/ scale of number line] x100. Spearman-Brown split half reliability coefficient for this sample was .917. This number range extends beyond curriculum expectations, which involve the knowledge and manipulation of numbers to at least 10 (Ministry of Education, 2013). However, ceiling performance of this sample on the numbers subtest of the Bracken School Readiness Test (not reported in this study) suggest that children in this sample know a number range that is above curriculum expectations.

Mathematical Achievement

Two mathematics subtests (Numerical Operations and Mathematical Reasoning) from the Wechsler Individual Achievement Test II (WIAT-II; Wechsler, 2005) were used. The Numerical Operations subtest assesses number reading and writing, counting and written calculation skills. The Mathematical Reasoning subtest assesses the application of numerical and mathematical concepts in practical situations and the ability to solve verbally presented word problems aided by visual cues. Performance was indexed by correct responses, and testing was discontinued after six consecutive errors (standard procedure). A composite score derived from adding the raw scores from each subtest was calculated. Raw scores are used as there is no standardisation of this assessment battery with a Singaporean sample.

Results

Estimation Patterns in the Number Line Estimation Task

In line with the first aim of the study, we fitted children's individual estimation patterns to the logarithmic, linear, unbounded power model, one-cycle power model and two-cycle power models (see Appendix) following the procedures of Slusser et al. (2013). We did not analyse the estimation patterns when the estimates failed to show a significant correlation with the target numbers (p > .05 for Spearman's correlation) or when 90% of the estimates were clustered in 10% or less of the number range, as suggested by Slusser et al. (2013). As this was the case for 48 (14.8%) participants, we report on a final sample of 276 participants (142 boys, M_{age} = 6.2 years, SD_{age} = 0.3 years). Models were compared in two stages: within theoretical account and between theoretical accounts, based on their Akaike Information Criterion adjusted for small samples (AICc). The model with the lowest AICc was chosen as the model that best explained the estimation patterns. We chose AICc as it accounts for model fit and the number of parameters included in each model, therefore allowing for comparisons across models with different number of parameters (Burnham & Anderson, 2004).

Descriptive statistics are presented in Table 1. This sample's accuracy (PAE) on the number line task was similar to that found in previous studies with children in a similar age range (Laski & Siegler, 2007; Siegler & Booth, 2004; Slusser et al., 2013).

Table 1

Descriptive Statistics

Variable	*M*	*SD*	Min.	Max.	Skewness	Kurtosis
Mathematical achievement	28.82	5.46	14.00	41.00	-0.28	2.63
Number line estimation (Percent absolute error)	16.15	6.67	4.67	37.89	0.47	2.77

Note. N = 276. For number line estimation smaller values represent better performance.

Comparing Within Theoretical Accounts

Within the representational shift account estimation patterns were less frequently best explained by a logarithmic (35.9%, M_{AICc} = 158.19) compared to a linear model (64.1%, M_{AICc} = 145.58; see top panel of Supplementary Table 1). These results suggest that almost two thirds of children have developed a linear representation of magnitudes up to 100, while a third of them still have a logarithmic representation.

https://doi.org/10.5964/jnc.10557

Within the proportional judgment account, 57.6% (M_{AICc} = 148.54) of estimation patterns were best explained by an unbounded power model, 29.7% (M_{AICc} = 149.84) by a one-cycle power model, and 12.7% (M_{AICc} = 146.48) by a two-cycle power model (see top panel of Supplementary Table 2). These results suggest that over half of children are only able to use the start point of the number line to guide their estimates, over a quarter can use both start and end point, and over a tenth can also infer a midpoint.

Considering that AICc is a relative measure, and that in some cases we had noticed very small AICc differences between models, it was necessary to account for the strength of the evidence for each model. Burnham and Anderson (2004) proposed that if the difference in AICc values (ΔAICc) between any given model and the one that best explains the data is less than 2, then the alternative model has substantial support. Therefore, we conducted all analyses again excluding estimation patterns for which the ΔAICc between the two best models was less than 2.

Within the representational shift account, 10.9% of estimation patterns showed similar support for both models (i.e, ΔAICc < 2), 29.7% (M_{AICc} = 156.46) showed stronger support for the logarithmic model, while 59.4% (M_{AICc} = 144.48) showed stronger support for the linear model (see bottom panel of Supplementary Table 1). Within the proportional judgment account, we found that 26.1% of the estimation patterns showed similar support for at least two models, while 47.1% (M_{AICc} = 148.62) showed stronger support for the unbounded power model, 17.4% (M_{AICc} = 142.27) for the one-cycle power model and 9.4% (M_{AICc} = 141.18) for the two-cycle power model (see bottom panel of Supplementary Table 2). Though overall trends remain consistent with the original analyses, these findings highlight that for a tenth and over a quarter of estimation patterns, for the representational shift and proportional judgment accounts respectively, using the models that best explain the estimation patterns was not a reliable index of performance.

Comparing Between Theoretical Accounts

We first consider the distribution across both accounts based on the models arising from the previous analyses, which we represent in a two-way table (see Table 2). We observe that the percentage of estimation patterns best explained by the logarithmic model decreases as the power models become more advanced, while this trend is opposite for the linear model. A Chi square test of independence (conducted in the R package Stats, R Core Team, 2019) showed a significant association between the best fit models of the representational shift and proportion judgment accounts, $\chi^2(2, 276) = 42.19$, $p < .001$. Estimation patterns best explained by the logarithmic model were overrepresented within those best explained by the unbounded power model (z = 6.1) and underrepresented within those best explained by the one (z = -3.1) and two-cycle power models (z = -4.7). The opposite was found for the estimation patterns best explained by the linear model. This suggests that more advanced models within the proportional judgment account are more consistent with a linear pattern, while less advanced ones are more consistent with a logarithmic pattern.

Table 2

Distribution of Models Across Both Theoretical Accounts

Best fitting model	**Logarithmic (*n* = 99)**	**Linear (*n* = 177)**
Unbounded (*n* = 159)	51% (81)	49% (78)
One-cycle (*n* = 82)	22% (18)	78% (64)
Two-cycle (*n* = 35)	0% (0)	100% (35)

Note. Two-way table representing the distribution of models that best explained estimation patterns across both theoretical accounts based on each theoretical account being considered independently. Models on the first row are based on the representational shift account, while models on the first column are based on the proportional judgment account. Percentages were obtained by dividing the counts of each cell by its corresponding row's total.

Notwithstanding, estimation patterns best explained by the unbounded power model were almost equally divided (n = 81 and 78) between the logarithmic and linear models when considered from a representational shift account. If we consider both theoretical accounts simultaneously, these results suggest that children who only use the start point of the number line to guide their estimates could show either a logarithmic or a linear representation of magnitude.

To extend this understanding we directly contrasted the models arising from both theoretical accounts (see top panel in Supplementary Table 3). Overall, we found that 60.5% of estimation patterns were best explained by models arising from the proportional judgment account (25.0% by the unbounded (M_{AICc} = 140.11), 23.2% by the one-cycle (M_{AICc} = 149.45), and 12.3% by the two-cycle power models (M_{AICc} = 147.70)). The remaining 39.5% were best explained by the models arising from the representational shift account (23.9% by a logarithmic model (M_{AICc} = 160.18), and 15.6% by a linear model (M_{AICc} = 138.53)). See Supplementary Table 4 for the overlap between these findings and those arising from the comparisons conducted within theoretical accounts.

As before, we also accounted for the strength of the evidence for each model by excluding estimation patterns for which the ΔAICc was less than two (see bottom panel of Supplementary Table 3). We found that 51.5% of estimation patterns showed similar support for at least two models, while 30.4% of estimation patterns showed stronger support for one of the models arising from the proportional judgment account (13.4% for the unbounded power model (M_{AICc} = 133.78), 8.7% for the one-cycle power model (M_{AICc} = 146.19), and 8.3% for the two-cycle power model (M_{AICc} = 141.22)). The remaining 18.1% showed stronger support for a model arising from the representational shift account (13.0% for the logarithmic model (M_{AICc} = 153.13), and 5.1% for the linear model (M_{AICc} = 129.02)). Table 3 shows all alternative models that had substantial support in each case. These results highlight the challenges involved in drawing conclusions from model comparisons when multiple models explain the data equally well. We will return to this point in the Discussion.

Table 3

Number of Individual Estimation Patterns in the 0-100 Number Line for Which an Alternative Model Showed Substantial Support

	Model that best explains the estimation patterns				
Alternative model with similar support	**Logarithmic model (*n* = 66)**	**Linear model (*n* = 43)**	**Unbounded power model (*n* = 69)**	**One-cycle power model (*n* = 64)**	**Two cycle power model (*n* = 34)**
ΔAICc < 2 with linear model	–	0	9	8	0
ΔAICc < 2 with logarithmic model	2	–	18	21	8
ΔAICc < 2 with unbounded power model	30	22	–	20	3
ΔAICc < 2 with one-cycle power model	6	12	10	–	7
ΔAICc < 2 with two-cycle power model	0	3	4	3	–

The Relationship Between Number Line Estimation Patterns and Mathematical Performance

To address the second aim of this study we created groups based on the results obtained from the first aim and then compared those groups on their accuracy on mathematical achievement. First, we compared models within theoretical accounts, and then between theoretical accounts. In both cases, analyses were run first including all estimation patterns, and then excluding those that had shown similar support for more than one model (i.e., ΔAICc < 2). Non-parametric tests were used as assumptions of normality were not met. For comparisons within the representational shift account the Wilcoxon rank sum test was used. In all other cases, the Kruskal Wallis test was used, followed up by Dunn's test with a Holm correction. These analyses were conducted using the Stats (R Core Team, 2019) and RStatix (Kassambara, 2020) packages in R (R Core Team, 2019). Descriptive statistics of mathematical achievement for each group can be found in Supplementary Tables 5 through 7.

Comparing Within Theoretical Accounts

Within the representational shift account, the linear group outperformed the logarithmic group in mathematical achievement, W = 11950, $p < .001$, r = .302. These results remained consistent, W = 9609, $p < .001$, r = .350, when excluding cases that showed similar support for both models. Within the proportional judgment account, we found evidence for significant group differences, $H(2)$ = 10.45, p = .005, η^2 = .031, in mathematical achievement. Follow-up multiple comparisons showed significant differences in favour of the two-cycle power model when compared to the unbounded (z = 3.19, p = .004), and the one-cycle power model (z = 2.75, p = .012). No significant differences were found between the unbounded and the one-cycle power model (z = .308, p = .758). Group differences remained significant, $H(2) = 10.22$, p = .006, $\eta^2 = .041$, when excluding cases that showed similar support for more than one model. However, follow-up multiple comparisons only showed significant differences between the two-cycle and the unbounded power model (z = 3.18, p = .004). The one-cycle was not significantly different from the unbounded (z = .976, p = .329), nor than the two-cycle power model group (z = 2.13, p = .066).

Comparing Between Theoretical Accounts

When considering all models simultaneously, we found significant differences in mathematical achievement, $H(4)$ = 41.90, $p < .001$, $\eta^2 = .140$. Follow-up multiple comparisons showed that the logarithmic group was outperformed by the linear (z = 3.84, p = .001), unbounded (z = 5.55, $p < .001$), one-cycle (z = 3.34, p = .006), and two-cycle power model groups (z = 5.23, $p < .001$). All other comparisons were not statistically significant ($p > .05$). Group differences remained significant, $H(4)$ = 33.71, p < . 001, $\eta^2 = .230$, when excluding cases that showed similar support for more than one model. Follow-up multiple comparisons showed that the logarithmic group was outperformed by the linear (z = 3.50, p = .004), unbounded (z = 4.52, $p < .001$), and two-cycle power model groups (z = 4.91, $p < .001$). However, differences between the logarithmic and one-cycle power model group were not significant (z = 2.05, p = .244). All other comparisons remained not significant ($p > .05$).

Discussion

Number Line Estimation Patterns

The first aim of this study was to determine which mathematical functions best explain children's estimation patterns on the 0-100 number line task. Specifically, we examined estimation patterns based on two dominant perspectives – logarithmic to linear representational shift and proportional judgment. We first examined them separately, and then compared them directly.

Based on the representational shift account (Siegler & Opfer, 2003), it is anticipated that children's estimation patterns change from logarithmic to linear with increased experience within a number range. The shift from logarithmic to linear seems to have occurred earlier in our sample compared to previous studies. For example, Laski and Siegler (2007) found that 43% of kindergarteners (M_{age} = 6.10 years) and 53% of first graders' (M_{age} = 7.19 years) estimation patterns on the 0-100 number line were best fit by a linear model, compared to 64% (59.4% after excluding the 10% of estimates that showed similar support for both models) in the current sample (M_{age} = 6.2 years). Studies with younger children have found the linear model to become predominant in second grade, with the logarithmic model prevailing during kindergarten (M_{age} = 5.7 years) and first grade (Praet & Desoete, 2014; Siegler & Booth, 2004). This earlier shift to a linear representation aligns with previous findings that suggest an earlier development of some mathematical skills in Singaporean children compared to some of their international counterparts (Yao et al., 2017), and highlights the role of education in the performance on the number line estimation task.

Based on the proportional judgment account (Barth & Paladino, 2011), we anticipate that with time children reduce their estimation bias and increase the number of benchmarks they use. Most children in our sample tend to use only the start point of the number line to guide their estimates, with fewer children being able to use both start and endpoint, and a minority also capable of inferring a midpoint. At first glance, these findings seem appropriate given the young age of children in this sample who have not started primary school. However, our sample showed a higher percentage of children's estimation patterns being best explained by the unbounded power model (57.6%) compared to that found in previous studies with a similar age group (47% in Sasanguie et al., 2016, and 30% in Slusser & Barth, 2017 and Xing et al., 2021), although our findings resemble those of Sasanguie et al. (2016) when excluding the 26% of participants who showed similar support for more than one model. In addition, in this sample almost 10% of children whose estimates are best explained by an unbounded power model show the more unusual under estimation pattern ($\beta > 1$). This seems to suggest an overrepresentation of this phenomenon in our sample (see Xing et al., 2021), though possible comparisons with previous studies are very limited due to lack of reporting.

Comparing models across theoretical perspectives might be controversial, and it is not without its difficulties (see Opfer et al., 2011 for similarities between the logarithmic and unbounded power model, and between the cyclical power models and the linear model). However, considering models from both theoretical perspectives simultaneously seemed essential to better understand what estimation patterns can tell us about children's performance on the number line task, and how that might change based on which theoretical viewpoint one stands from.

When looking at the distribution of models across both theoretical accounts, we found that the percentage of estimation patterns best explained by the linear model increases as the power models become more advanced (i.e., the percentage of children's estimates best explained by the linear and two-cycle power model is higher than the one best explained by the linear and one-cycle model, which is higher than the one best explained by the linear and unbounded power models). This trend was opposite for the estimation patterns best explained by the logarithmic model. This finding seems consistent with the idea that the use of additional benchmarks (i.e., more advanced models within the proportional judgement account) is more likely to correspond to more linear estimation patterns within the representational shift account. However, an unexpected finding was that approximately half of the estimation patterns best explained by the unbounded power model were also best explained by the linear model, while the other half was best explained by the logarithmic model.

These findings are unexpected, as we would not anticipate that children who only use the beginning of the number line to guide their estimates (i.e., unbounded power model) can achieve linear estimations. However, there are precedents of estimation patterns being similarly explained by unbounded and linear models. For example, in Link et al.'s (2014) study with adults, 47% of estimation patterns were best explained by an unbounded power model even though these were very precise estimates (i.e., almost linear). Furthermore, in Friso-van den Bos et al. (2015), the percentage of children's estimation patterns for whom the linear model represented the best fit after excluding the unbounded power model increased with age. The linear model became a more likely substitution of the unbounded power model over the logarithmic model by the time participants were in second grade. Link et al. (2014) in the context of their study suggest that the high level of accuracy in their participants' estimates meant that the unbounded power model would closely resemble a linear model. These findings suggest that it is possible that very accurate estimates (which resemble a linear model), can be better explained by an unbounded power model. Participants in our sample were not accurate enough to assume that the same phenomenon as in Link et al.'s study is in play here. But several characteristics of our sample's performance need to be considered when interpreting the concurrence between the unbounded and linear models.

First, it is important to note as shown earlier, that other models, beyond the ones that best explain the data might show similar support. For example, from the 78 participants for whom the unbounded power model and the linear model concurred, over 25% showed similar support between the unbounded and one of the other power models, while this was the case for 6% that showed similar support for the logarithmic and linear models.

In addition, almost 20% of the participants for whom the unbounded power model and the linear model concurred showed a less frequent pattern of $\beta > 1$, while all participants for whom the unbounded power model and the logarithmic model concurred showed the more frequent pattern $\beta < 1$. This might mean that, even though best explained by the unbounded power model, some of these children could be engaging with the task in a different way than other children whose estimates are also best explained by the unbounded power model.

Furthermore, from the estimation patterns in which the linear and unbounded power model had provided the best explanation within each theoretical account, the unbounded power model better explained 62.8% of them. Contrastingly, from the estimation patterns in which the unbounded power model and the logarithmic model had provided the best explanation within each theoretical account, the logarithmic model better explained 75.3% of them.

Finally, as mentioned earlier, estimation patterns best explained by one and two-cycle power models were still more likely to also be better explained by linear models, than by logarithmic models. In principle, it seems highly unlikely that a linear estimation pattern would result from the exclusive use of the origin of the number line, though some evidence of this possibility exists (e.g., Friso-van den Bos et al., 2015; Link et al., 2014). However, as proposed by Slusser et al. (2013), a variety of strategies might give rise to estimation patterns being best explained by an unbounded power model. That is, children might use the origin of the number line in different ways to guide their estimates (e.g., magnitude judgments that are open-ended, counting up from the origin by using arbitrary units). This possible variability makes it very difficult to speculate as to which strategies used by children might lead to a concurrence of these two models (i.e., linear and unbounded power model), and highlights the relevance of further studies focusing on the observation and analysis of how children go about solving the number line task.

From the comparison of all models to each other, we find that models arising from the proportional judgment perspective best explain most estimation patterns (60.5%, 30.4% after excluding estimation patterns with similar support for more than one model). When considering models individually, the logarithmic, unbounded power, and one-cycle power models account for slightly above 23 percent of children's estimation patterns each (this is around 13 percent for the logarithmic and unbounded power model, and 8 percent for the one-cycle power model after excluding estimation patterns showing similar support for more than one model), and the linear and two-cycle power model account for approximately 15 and 12 percent of children's estimation patterns respectively (8 and 5 percent after excluding estimation patterns with similar support for more than one model). These findings are consistent with previous studies in which best fitting models are mostly distributed between models arising from both theoretical accounts (e.g., Link et al., 2014; Slusser et al., 2013), though inconsistent with those that have found an overwhelming majority of estimation patterns being best explained by the models arising from a single theoretical account (e.g., Slusser & Barth, 2017; Xu et al., 2023; Yuan et al., 2020).

Most previous studies have limited their evaluation of estimation patterns to one theoretical account. This only allows the studies to examine whether or not performance is consistent with this account, but not if there is an alternative account that might be more consistent with performance patterns. Furthermore, most studies report the model with the lowest AIC as the 'best' model, ignoring the possibility that other models may explain the data equally well. Taken together our findings are hard to reconcile with the assumptions underlying the chosen theoretical accounts. For example, based on the representational shift account for 10% of participants, either model explained the data equally well. This seems incompatible with the idea that the shape of the estimation patterns can provide a direct reflection of children's mental representation of magnitude. However, this concurrent fit should be considered in light of overlapping waves theory (Siegler, 1996) which suggests that at any given point in time children have different representations and strategies, and might be used alternatively depending on the context.

Turning to the proportional judgment account, we found that over a quarter of estimation patterns could be explained equally well by more than one power model. In this regard, the unbounded, one-cycle and two-cycle power models, do not seem to be reliable indices of the number and types of benchmarks used by children when performing the number line estimation task. This does not imply that children do not use a variety of strategies (or benchmarks) to solve the number line task, but that, at least for our sample, the models arising from this theoretical account do not seem to be very informative regarding children's strategies. Furthermore, when conducting comparisons between theoretical accounts, we found that over a quarter of estimation patterns best explained by the unbounded power model had similar support for the linear model. Conversely, from the estimation patterns best explained by the linear model, over half show similar support for the unbounded power model. This concurrence between unbounded power and linear models puts into question whether the unbounded power model necessarily represents children who can only rely on the origin of the number line, as they are unlikely to produce linear estimates (though see Link et al., 2014).

Estimation Patterns and Mathematical Performance

The second aim of our study was to determine if the models that best explain children's estimation patterns were associated with their mathematical performance. When comparisons were conducted within theoretical accounts, children whose estimation patterns were best explained by the linear model had higher mathematical achievement than children whose estimation patterns were best explained by the logarithmic model. This remained the case when excluding children whose estimation patterns showed similar support for both models. This is consistent with the idea that a more linear representation of magnitude is associated with higher levels of concurrent mathematical competence (Praet & Desoete, 2014; Xu et al., 2023). Within the proportional judgment account, children whose estimation patterns were best explained by the two-cycle model outperformed children whose estimation patterns were best explained by both the unbounded and one-cycle power models. This is consistent with previous findings by Xing et al. (2021), although in our sample, only the difference between the two-cycle and the unbounded power model remained significant when excluding children whose estimation patterns showed similar support for more than one model.

Finally, when comparing between theoretical accounts, children whose estimation patterns were best explained by the logarithmic model showed lower mathematical achievement than children whose estimation patterns were best explained by all other models, although the difference with the one-cycle power model became non-significant when excluding children whose estimation patterns showed similar support for more than one model. Our findings are partially consistent with those of Friso-van den Bos et al. (2015) in which children whose estimation patterns were best fit by the logarithmic model, or the unbounded power model were outperformed by those whose estimation patterns were best fit by either the linear or one-cycle power model.

The finding that children whose estimation patterns were best explained by the unbounded power model outperformed those best explained by the logarithmic model is interesting for several reasons. If we consider both theoretical accounts, both models represent the least advanced stage (i.e., representation, strategy). Moreover, as mentioned prior, it has also been proposed that the unbounded power model under certain circumstances can resemble a logarithmic function (Opfer et al., 2011). This is also reinforced by the high frequency of children whose estimation patterns were best explained by the logarithmic model while showing very similar support for the unbounded power model, and vice versa. Therefore, such contrasting levels of mathematics achievement would not be expected. However, this also needs to be contextualised by considering that as a group, children whose estimation patterns were best explained by the unbounded power model produced more accurate estimates than children whose estimation patterns were best explained by the logarithmic model (M_{PAE} = 13.32 and 24.14, respectively; W = 4448.5, p < .001, r = .823). Given the known association between accuracy on the number line task and mathematical achievement (Schneider et al., 2018), it is possible that the differences observed in mathematical achievement between these two groups are at least partly explained by differences in number line accuracy.

Our findings, combined with those from other studies (e.g., Friso-van den Bos et al., 2015; Xing et al., 2021) suggest that there is a relationship between estimation patterns on the number line estimation task and mathematical achievement. Children who produce more advanced estimation patterns (within each theoretical account), tend to show higher mathematical achievement. What remains to be answered is to what extent does estimation accuracy explain these differences in mathematical achievement.

Limitations

This study is limited to the use of model fitting to distinguish between two theoretical accounts. However, this is not the only way in which these theoretical accounts can be studied or compared. In addition, we only consider the original models arising from the representational shift and proportional judgment accounts. Other models exist within each of these accounts: a mixed logarithmic-linear model (Opfer et al., 2016), and a mixed cyclical power model (Zax et al., 2019). These models, each within its own theoretical account, allow for the combination of the two pure models (i.e., linear and logarithmic, and the one and two-cycle power models). The aim of these models is to quantify to which degree any given estimation pattern resembles each of the initial pure models, under the understanding that a combination of both might provide a better explanation to the estimation patterns that any pure model on its own. Furthermore, other theoretical accounts have proposed other models such as the double/segmented linear model (Ebersbach et al., 2008, 2015; Moeller et al., 2009), and the subtraction bias cyclic model (Cohen & Sarnecka, 2014). We decided to only focus on the originally presented models because these are the most frequently reported in the literature, and because increasing the number of models leads to an increase in number of comparisons, with an accompanying loss of statistical power.

Additionally, selecting one model that best explains children's individual estimation patterns might be an artificial differentiation, as a significant percentage of children's estimation patterns also showed substantial support for at least one other model. We attempted to address this by running the same analyses after excluding these cases. However, this leads to a loss of statistical power, which requires caution for the interpretation of these findings.

A final limitation of this study is that the target numbers used for the number line task were not equally distributed across the number range, showing oversampling of the lower half of the number range. This might have favoured some models (i.e., logarithmic) over others (i.e., cyclical power models).

Conclusion

The current study showed that most children's estimation patterns can be explained by one of the mathematical functions proposed by the representational shift and proportional judgment accounts, and that which model best explains children's estimation patterns is associated with their mathematical achievement. However, this study highlighted the obstacles and limitations involved in using estimation patterns as an index of performance, especially given that for a high percentage of children's estimation patterns at least one other model explained the data equally well as the best model.

Funding: 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.

Acknowledgments: The authors have no additional (i.e., non-financial) support to report.

Competing Interests: The authors have declared that no competing interests exist.

Author Contributions: Carola Ruiz – Conceptualization, formal analysis, funding acquisition, methodology, writing (original draft, review and editing). **Saskia Kohnen** – Conceptualization, methodology, writing (review and editing). **Rebecca Bull** – Conceptualization, funding acquisition, methodology, writing (review and editing).

Data Availability: The dataset for the project "Number Line Estimation Patterns and Their Relationship with Mathematical Performance" is stored at the Macquarie University Research Data Repository (https://doi.org/10.25949/22558528.v1).
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).

Supplementary Materials

The Supplementary Materials contain the following items (for access see Index of Supplementary Materials below):

- The code used to conduct the analyses
- Descriptive statistics and additional analyses for the study

Index of Supplementary Materials

Ruiz, C., Kohnen, S., & Bull, R. (2023a). *Supplementary materials to "Number line estimation patterns and their relationship with mathematical performance"* [Code]. OSF. https://osf.io/jat5h/

Ruiz, C., Kohnen, S., & Bull, R. (2023b). *Supplementary materials to "Number line estimation patterns and their relationship with mathematical performance"* [Descriptive statistics and additional analyses]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.12695

References

Barth, H. C., & Paladino, A. M. (2011). The development of numerical estimation: Evidence against a representational shift. *Developmental Science, 14*(1), 125–135. https://doi.org/10.1111/j.1467-7687.2010.00962.x

Bouwmeester, S., & Verkoeijen, P. P. J. L. (2012). Multiple representations in number line estimation: A developmental shift or classes of representations? *Cognition and Instruction, 30*(3), 246–260. https://doi.org/10.1080/07370008.2012.689384

Bull, R., Lee, K., & Muñez, 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

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods Research, 33*(2), 261–304. https://doi.org/10.1177/0049124104268644

Cohen, D. J., & Sarnecka, B. W. (2014). Children's number-line estimation shows development of measurement skills (not number representations). *Developmental Psychology, 50*(6), 1640–1652. https://doi.org/10.1037/a0035901

Dackermann, T., Fischer, U., Huber, S., Nuerk, H. C., & Moeller, K. (2016). Training the equidistant principle of number line spacing. *Cognitive Processing, 17*(3), 243–258. https://doi.org/10.1007/s10339-016-0763-8

Dackermann, T., Huber, S., Bahnmueller, J., Nuerk, H.-C., & Moeller, K. (2015). An integration of competing accounts on children's number line estimation. *Frontiers in Psychology, 6*, Article 884. https://doi.org/10.3389/fpsyg.2015.00884

Department of Statistics Singapore. (2013). *Find data*. https://www.singstat.gov.sg/find-data

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), 1–17. https://doi.org/10.1016/j.jecp.2007.08.006

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

Elofsson, J., Gustafson, S., Samuelsson, J., & Träff, U. (2016). Playing number board games supports 5-year-old children's early mathematical development. *The Journal of Mathematical Behavior, 43*, 134–147. https://doi.org/10.1016/j.jmathb.2016.07.003

Friso-van den Bos, I., Kroesbergen, E. H., & Van Luit, J. E. H. (2018). Counting and number line trainings in kindergarten: Effects on arithmetic performance and number sense. *Frontiers in Psychology, 9*, Article 975. https://doi.org/10.3389/fpsyg.2018.00975

Friso-van den Bos, I., Kroesbergen, E. H., Van Luit, J. E. H., Xenidou-Dervou, I., Jonkman, L. M., Van der Schoot, M., & Van Lieshout, E. C. D. M. (2015). Longitudinal development of number line estimation and mathematics performance in primary school children. *Journal of Experimental Child Psychology, 134*, 12–29. https://doi.org/10.1016/j.jecp.2015.02.002

Hollands, J. G., & Dyre, B. P. (2000). Bias in proportion judgments: The cyclical power model. *Psychological Review, 107*(3), 500–524. https://doi.org/10.1037/0033-295X.107.3.500

Huber, S., Moeller, K., & Nuerk, H. C. (2014). Dissociating number line estimations from underlying numerical representations. *The Quarterly Journal of Experimental Psychology, 67*(5), 991–1003. https://doi.org/10.1080/17470218.2013.838974

Kassambara, A. (2020). *Rstatix: Pipe-friendly framework for basic statistical tests* [Computer software] (R Package Version 0.6.0).

Kim, D., & Opfer, J. E. (2017). A unified framework for bounded and unbounded numerical estimation. *Developmental Psychology, 53*(6), 1088–1097. https://doi.org/10.1037/dev0000305

Kucian, K., Grond, U., Rotzer, S., Henzi, B., Schönmann, C., Plangger, F., Gälli, M., Martin, E., & von Aster, M. (2011). Mental number line training in children with developmental dyscalculia. *NeuroImage, 57*(3), 782–795. https://doi.org/10.1016/j.neuroimage.2011.01.070

Laski, E. V., & Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Development, 78*(6), 1723–1743. https://doi.org/10.1111/j.1467-8624.2007.01087.x

Liang, Y., Zhang, L., Wang, C., & Liu, Y. (2021). Performance patterns and strategy use in number line estimation among preschool children with different spontaneous focusing on numerosity tendencies. *Infant and Child Development, 30*(1), Article e2203. https://doi.org/10.1002/icd.2203

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*, Article 1021. https://doi.org/10.3389/fpsyg.2013.01021

Link, T., Moeller, K., Huber, S., Fischer, U., & Nuerk, H. C. (2013). Walk the number line – An embodied training of numerical concepts. *Trends in Neuroscience and Education, 2*, 74–84. https://doi.org/10.1016/j.tine.2013.06.005

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

Maertens, B., De Smedt, B., Sasanguie, D., Elen, J., & Reynvoet, B. (2016). Enhancing arithmetic in pre-schoolers with comparison or number line estimation training: Does it matter? *Learning and Instruction, 46*, 1–11. https://doi.org/10.1016/j.learninstruc.2016.08.004

Ministry of Education. (2013). *Nurturing early learners: A curriculum for kindergartens in Singapore.* Numeracy. https://www.ecda.gov.sg/growatbeanstalk/Documents/MOE%20NEL%20Resources/NEL_educators%20guide%20vol%206_numeracy.pdf

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

Opfer, J. E., & Siegler, R. S. (2007). Representational change and children's numerical estimation. *Cognitive Psychology, 55*, 169–195. https://doi.org/10.1016/j.cogpsych.2006.09.002

Opfer, J. E., Siegler, R. S., & Young, C. J. (2011). The powers of noise-fitting: Reply to Barth and Paladino. *Developmental Science, 14*(5), 1194–1204. https://doi.org/10.1111/j.1467-7687.2011.01070.x

Opfer, J. E., Thompson, C. A., & Kim, D. (2016). Free versus anchored numerical estimation: A unified approach. *Cognition, 149*, 11–17. https://doi.org/10.1016/j.cognition.2015.11.015

Peeters, D., Verschaffel, L., & Luwel, K. (2017). Benchmark-based strategies in whole number line estimation. *British Journal of Psychology, 108*, 668–686. https://doi.org/10.1111/bjop.12233

Praet, M., & Desoete, A. (2014). Number line estimation from kindergarten to grade 2: A longitudinal study. *Learning and Instruction, 33*, 19–28. https://doi.org/10.1016/j.learninstruc.2014.02.003

Ramani, G. B., & Siegler, R. S. (2011). Reducing the gap in numerical knowledge between low- and middle-income preschoolers. *Journal of Applied Developmental Psychology, 32*(3), 146–159. https://doi.org/10.1016/j.appdev.2011.02.005

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Sasanguie, D., Verschaffel, L., Reynvoet, B., & Luwel, K. (2016). The development of symbolic and non-symbolic number line estimations: Three developmental accounts contrasted within cross-sectional and longitudinal data. *Psychologica Belgica, 56*(4), 382–405. https://doi.org/10.5334/pb.276

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

Siegler, R. (1996) *Emerging minds: The process of change in children's thinking*. Oxford University Press.

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., & Booth, J. L. (2005). Development of numerical estimation. A review. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 197-212). Psychology Press.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14*(3), 237–243. https://doi.org/10.1111/1467-9280.02438

Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games-but not circular ones-improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology, 101*(3), 545–560. https://doi.org/10.1037/a0014239

Slusser, E. B., & Barth, H. (2017). Intuitive proportion judgment in number-line estimation: Converging evidence from multiple tasks. *Journal of Experimental Child Psychology, 162*, 181–198. https://doi.org/10.1016/j.jecp.2017.04.010

Slusser, E. B., Santiago, R. T., & Barth, H. (2013). Developmental change in numerical estimation. *Journal of Experimental Psychology: General, 142*(1), 193–208. https://doi.org/10.1037/a0028560

Stevens, S. S. (1975). *Psychophysics. Introduction to its perceptual, neural, and social prospects.* John Wiley & Sons.

Thompson, C. A., Fitzsimmons, C. J., Mielicki, M., Scheibe, D. A., Sidney, P. G., Schiller, L., Siegler, R., & Opfer, J. E. (2022). Comments regarding numerical estimation strategies are correlated with math ability in school-age children. *Cognitive Development, 62*, Article 101188. https://doi.org/10.1016/j.cogdev.2022.101188

Wechsler, D. (2005). *Wechsler Individual Achievement Test 2nd edition* (WIAT II). The Psychological Corporation.

Whyte, J. C., & Bull, R. (2008). Number games, magnitude representation, and basic number skills in preschoolers. *Developmental Psychology, 44*(2), 588–596. https://doi.org/10.1037/0012-1649.44.2.588

Xing, C., Zax, A., George, E., Taggart, J., Bass, I., & Barth, H. (2021). Numerical estimation strategies are correlated with math ability in school-aged children. *Cognitive Development, 60*, Article 101089. https://doi.org/10.1016/j.cogdev.2021.101089

Xu, X., Chen, C., Pan, M., & Li, N. (2013). Development of numerical estimation in Chinese preschool children. *Journal of Experimental Child Psychology, 116*(2), 351–366. https://doi.org/10.1016/j.jecp.2013.06.009

Xu, X., Chen, C., Wang, L., Zhao, M., Xin, Z., & Liu, H. (2023). Longitudinal relationship between number line estimation and other mathematical abilities in Chinese preschool children. *Journal of Experimental Child Psychology, 228*, Article 105619. https://doi.org/10.1016/j.jecp.2022.105619

Yao, S.-Y., Muñez, 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

Zax, A., Slusser, E., & Barth, H. (2019). Spontaneous partitioning and proportion estimation in children's numerical judgments. *Journal of Experimental Child Psychology, 185*, 71–94. https://doi.org/10.1016/j.jecp.2019.04.004

Appendix

Formulas of models to be fitted to the number line data based on Slusser et al. (2013) and examples

(1) **Constrained Linear Model**

$y = ax + b,$

$b \geq 0$

Initial parameter values (slope = 1; intercept = 0)

(2) **Constrained Logarithmic Model**

$y = a(\ln(x)) + b$

$b \geq 0$

Initial parameter values (slope = 1; intercept = 0)

(3) **Unbounded Power Model**

$y = a(x^b)$

Initial parameter values (β = 1; range = 10)

(4) **One-Cycle Power Model**

$$y = \left(\frac{x^b}{x^b + (r-x)^b}\right) \times r$$

$r = constant$

Initial parameter values (β = 1; range = 100)

(5) **Two-Cycle Power Model**

For $x < \frac{r}{2}$:

$$y = r \times \left(0.5 \times \left(\frac{x^b}{x^b + \left(\frac{r}{2} - x\right)^b}\right)\right)$$

Otherwise:

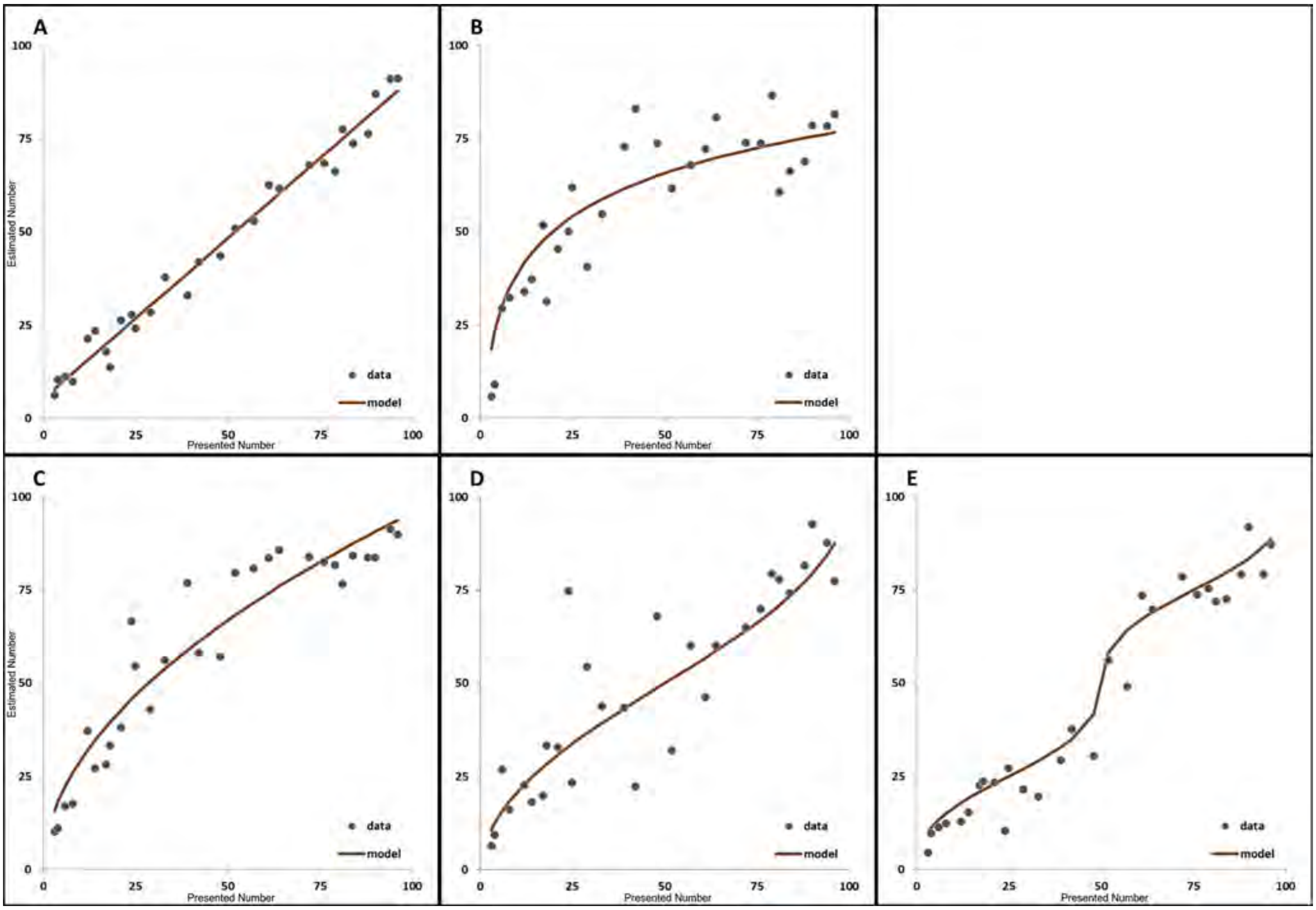
$$y = r \times \left(0.5 \times \left(\frac{\left(x - \frac{r}{2}\right)^b}{\left(x - \frac{r}{2}\right)^b + (r-x)^b}\right) + 0.5\right)$$

$r = constant$

Initial parameter values (β = 1; range = 100)

Figure A.1

Examples of Estimation Patterns Best Explained by Each Model in the Number Line Task



Note. Examples of estimation patterns in the 0-100 number line task best explained by a linear model (A), logarithmic model (B), unbounded power model (β < 1, C), one-cycle power model (β < 1, D) and two-cycle power model (β < 1, E). Representations of power models differ based on the β value being below or above 1.

MCLS

Journal of Numerical Cognition (JNC) is an official journal of the Mathematical Cognition and Learning Society (MCLS).

leibniz-psychology.org

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology (ZPID), Germany.