PAPER

123s and ABCs: developmental shifts in logarithmic-to-linear responding reflect fluency with sequence values

Michelle Hurst, K. Leigh Monahan, Elizabeth Heller and Sara Cordes

Department of Psychology, Boston College, USA

Abstract

*When placing numbers along a number line with endpoints 0 and 1000, children generally space numbers logarithmically until around the age of 7, when they shift to a predominantly linear pattern of responding. This developmental shift of responding on the number placement task has been argued to be indicative of a shift in the format of the underlying representation of number (Siegler & Opfer, 2003). In the current study, we provide evidence from both child and adult participants to suggest that performance on the number placement task may not reflect the structure of the mental number line, but instead is a function of the fluency (i.e. ease) with which the individual can work with the values in the sequence. In Experiment 1, adult participants respond logarithmically when placing numbers on a line with less familiar anchors (1639 to 2897), despite linear responding on control tasks with standard anchors involving a similar range (0 to 1287) and a similar numerical magnitude (2000 to 3000). In Experiment 2, we show a similar developmental shift in childhood from logarithmic to linear responding for a non-numerical sequence with no inherent magnitude (the alphabet). In conclusion, we argue that the developmental trend towards linear behavior on the number line task is a product of successful strategy use and mental fluency with the values of the sequence, resulting from familiarity with endpoints and increased knowledge about general ordering principles of the sequence.
A video abstract of this article can be viewed at: http://www.youtube.com/watch?v=zg5Q2LIFk3M*

Introduction

The ability to compare and estimate symbolic (e.g. Arabic numerals) and non-symbolic (e.g. an array of dots) representations of number has been shown to strongly predict arithmetic and mathematical ability (Halberda, Mazzoco & Feigenson, 2008; Geary, Bailey & Hoard, 2009; Siegler & Booth, 2004). Much research has focused on important questions such as how our understandings of both symbolic and non-symbolic number develop throughout the lifespan with increased experience and education (e.g. Halberda & Feigenson, 2008; Siegler & Opfer, 2003). Moreover, a significant amount of work has been dedicated to understanding exactly how number is represented in the mind, specifically investigating the format of the mapping between objective number (e.g. the actual number of items within an array or the value specified by an Arabic numeral) and subjective representations of number (i.e. the *perceived* size of these values;

Brannon, Wusthoff, Gallistel & Gibbon, 2001; Moyer & Landauer, 1967; Siegler & Opfer, 2003; Whalen, Gallistel & Gelman, 1999).

The question of how an objective stimulus is mapped to a subjective representation is neither new nor unique to the field of numerical cognition. Historical laws of sensation, such as Weber's law and the Weber-Fechner law, have dominated theories of sensation and perception for over a century. These classic psychophysical investigations have been extended to models of subjective representation of number and quantity leading to two prominent theories suggesting that the mapping between objective and subjective number is either: (1) logarithmic, with constant variability across magnitudes (Logarithmic Ruler model; Dehaene, 1992, 2003) or (2) linear, with increasing variability with magnitude (Scalar Variability model; Gibbon, 1977; Gibbon & Church, 1981). According to the Logarithmic Ruler model, larger numbers are subjectively closer together than smaller

Address for correspondence: Michelle Hurst, Boston College Department of Psychology, 140 Commonwealth Ave., 300 McGuinn Hall, Chestnut Hill, MA 02467, USA; e-mail: hurstm@bc.edu

numbers, but each value has the same absolute level of imprecision in their individual representation. Alternatively, the Scalar Variability model posits a linear scaling of number such that all numbers are equally spaced on the mental number line, but larger numbers have a more variable representation than smaller numbers.

Despite opposing models, behavioral investigations have found it difficult to distinguish between these two accounts of number representation. Tasks requiring estimation and comparisons of number demonstrate that larger numbers are treated with less absolute precision, as explained by Weber's law (Buckley & Gillman, 1974; Halberda & Feigenson, 2008; Jordan & Brannon, 2006). Yet this adherence to Weber's law fails to distinguish between the linear and the logarithmic ruler theories. The imprecision of large numbers during these tasks could be due to either the condensed nature of the large end of the spectrum (Logarithmic Ruler model) or due to the extra noise in large number representation (Scalar Variability model). Several studies have attempted to separate these theories, but results have been inconsistent. Behavioral tests of nonverbal arithmetic in animals and humans have provided support for a scalar variability account (Brannon *et al.*, 2001; Cordes, Gallistel, Gelman & Latham, 2007a; Cordes, King & Gallistel, 2007b; Gibbon & Church, 1981, 1990; but see Dehaene, 1992, 2001, 2003; Gallistel, Brannon, Gibbon & Wusthoff, 2001). On the other hand, neurological evidence from 'number neurons' in monkeys has been suggestive of a logarithmic ruler (Neider & Miller, 2003).

Recently, Siegler and Opfer (2003) have provided persuasive evidence from their number placement (NP) task suggesting that both formats of number representation may be simultaneously present and operate in tandem across development. In their NP task, participants indicated where they thought a number fell on a line between two given numerical endpoints. Results revealed that young children initially respond in a logarithmic fashion on a 0 to 1000 number line, spacing smaller values further apart than larger values; however, sometime between 2nd and 4th grade, their number placements become increasingly linear, with equal spacing between values. Importantly, the authors posit that this pattern of findings is reflective of a developmental shift in the underlying representation of number, stating that the task 'provides particularly direct information about [children's] representations of numerical magnitude' (Siegler & Booth, 2005, p. 207). Specifically, they argue that early in development children primarily represent number via a logarithmically scaled representation of number, but education and experience give way to the development of a primarily linearly scaled representation (Opfer & Siegler, 2007; Opfer & Thompson, 2008; Siegler & Booth, 2004; Siegler & Opfer, 2003). Siegler, Opfer and colleagues suggest that both types of number mapping are simultaneously present, with the dominant mapping changing over the course of development.

Since Siegler and Opfer (2003), researchers have extended the use of the NP task to assess numerical representation in multiple contexts and its relationship with other estimation and mathematical tasks. Studies have revealed that NP task performance is predictive of mathematical competence (Booth & Siegler, 2008) and correlated with performance on other numerical estimation and magnitude comparison tasks (Laski & Siegler, 2007). Work investigating representations of number across cultures suggests that Amazonian indigene groups who lack distinct words for large numbers and who have very little, if any, formal education respond logarithmically (Dehaene, Izard, Spelke & Pica, 2008). In addition, the task has been extended to examine children's developing sense of fractions, decimals, and salaries (Opfer & DeVries 2008; Schneider, Grabner & Paetsch, 2009; Siegler, Thompson & Schneider, 2011) while also providing a concrete frame of reference for the development of educational interventions (Siegler & Ramani, 2008, 2009). Thus, the NP task has fostered diverse, fruitful, and important areas of research, proving it to be an important addition to our repertoire of numerical tasks.

However, there is continued disagreement about whether data from NP tasks provide insight into the structure of the internal representation of number (Barth & Paladino, 2011; Cantlon, Cordes, Libertus & Brannon, 2009; Cohen & Blanc-Goldhammer, 2011; Slusser, Santiago & Barth, 2013). Barth and colleagues (e.g. Barth & Paladino, 2011; Barth, Slusser, Cohen & Paladino, 2011; Slusser *et al.*, 2013) have systematically compared various models of NP data from both children and adults. Importantly, accurate performance on the NP task requires the ability to make a proportional judgment by comparing the given number to both the small and large endpoints, making it a proportion judgment task. For example, in order to accurately (i.e. linearly) place 20 on a number line with endpoints of 0 and 100, a child must compare the relative distance between 0 and 20 and between 20 and 100 and understand that 20 is closer to 0 than to 100. Consistent with this account, people spontaneously spend a significant amount of time looking at the center of the line when engaging in NP tasks (Sullivan, Juhasz, Slattery & Barth, 2011), suggesting that a reasonable strategy to achieve linearity is to create anchors or landmarks throughout the line, such as the halfway point. Thus, a

power function, which takes into account the individual's use of anchor points along the line, has been argued to better explain the behavioral evidence both mathematically (Slusser *et al.*, 2013) and theoretically (Hollands & Dyre, 2000). On this account, performance on the NP task is dependent upon the individual's understanding of proportion and not indicative of the format of the underlying subjective representation of number.

If NP data do not reflect the underlying representation of number, then what contexts and/or task-related factors determine when a child will perform linearly or logarithmically? Here, we argue for two factors that influence response patterns. First, we contend that the ability to implicitly compute proportions across the range of values presented is predictive of linear responding. By 'implicitly computing proportions', we refer to an approximate understanding of the number of values that fall between the two endpoints, as well as an ability to quickly and easily identify standard anchors within the range (e.g. the midpoint). The identification of intermediate anchors is reasonably easy to do when presented standard (i.e. ending in '0') endpoints (e.g. an adult knows half of 1000 is 500). With non-standard endpoints (i.e. non-decade numerals), however, online proportion computations become more difficult (e.g. what is the half-way point between 1639 and 2897?) and a more child-like, logarithmic pattern of responding would likely result. Thus, it follows that the ease with which the endpoints can be mentally manipulated and intermediate anchors can be determined, rather than the particular numerical magnitude of the values in question, will predict linear performance on the NP task. We explore this hypothesis in Experiment 1.

Second, we posit that a strong understanding of the relative ordering of *all* of the values in the range (e.g. having comparable knowledge of the ordering of values from 0 to 50 as those from 50 to 100) supports linear responding. If only partial fluency is attained, then greater weight will likely be placed upon the portion of the range that has been mastered, resulting in a greater proportion of the line being dedicated to the placement of these values. Consequently, mastered values will have greater spacing between values, leaving less space for the placement of other non-fluent values, resulting in the compression of less familiar values (Lourenco & Longo, 2009). Given that ordered lists are learned and practiced in order (such that children learn 1-2-3 *before* learning 8-9-10), partial fluency likely results in a greater emphasis on values at the beginning of the list, resulting in a logarithmic pattern of responding (greater spacing between earlier values, with a greater portion of the number line dedicated to smaller values). If it is the case

that linear responding is dependent upon fluency with the relative ordering (and not the relative magnitude) of the values in the range, then a logarithmic-to-linear developmental shift may be observed on placement tasks involving numerical *and* non-numerical learned sequences. Experiment 2 provides a direct test of this hypothesis.

In our first experiment, we presented adult subjects with an NP task involving non-standard endpoints (numerical values that do not end in 0 and thus are less easily manipulated). We compared their logarithmic performance on the non-standard endpoint task to their linear performance on two NP tasks involving similar numerical magnitudes or the same numerical range. Critically, both of these control tasks had at least one standard endpoint (ending in 0). In our second experiment, we report a similar logarithmic-to-linear developmental shift in responding on a non-numerical placement task, an alphabet line task, in which ordering cannot be dependent upon underlying magnitudes. Together these results strongly suggest that NP data do not reflect the format of the underlying representation of number but are instead indicative of fluency with the values within the sequence.

Experiment 1

Siegler and colleagues report that children may respond in a linear fashion with smaller numerical ranges (e.g. 0 to 10), but still respond logarithmically for larger numerical ranges (e.g. 0 to 100; Siegler & Booth, 2004; Siegler & Opfer, 2003; Siegler, Thompson & Opfer, 2009). Furthermore, when children place an identical subset of numbers on a 0 to 10 number line and on a 0 to 100 number line, the relative placement (and thus logarithmic or linear pattern of responding) of these exact same numbers varies depending on the context (i.e. the range given; Siegler & Opfer, 2003). This has been taken as evidence to support the simultaneous existence of logarithmic and linear representations, as proposed by Siegler and Opfer (2003). In Experiment 1a, we explore this context effect with adult subjects by examining whether it is the absolute magnitude of the endpoints (e.g. 100 is a larger number than 10), the magnitude of the range of values specified by the endpoints (e.g. 0 to 100 includes more values than 0 to 10), or alternatively, familiarity with the actual endpoint values (e.g. children are more experienced in working with 10 than with 100) that results in logarithmic responding. In Experiment 1b, we replicate these findings with a different group of adult subjects, while at the same time correcting for an error in the distribution of values tested.

Experiment 1a

Method

Participants

Participants were 40 undergraduate students (18 male; M = 19.5 years; Range: 18 to 23 years) from Boston College, who completed all tasks for course credit.

Experimental design and materials

Participants completed four placement tasks (adapted from Siegler & Opfer, 2003) within a single experimental session, in a quiet room in the laboratory. For each task, a 25 cm line was presented on a computer screen with specific endpoints dependent on the task. On each trial, a test value was presented 2 cm above the center of the line and participants were given 3 seconds to estimate the position of the test value on the line by using the mouse to move a hash mark to the desired position and clicking to select their answer. Within each task, test values were presented in a random order. After each trial, participants had to click a 'Push for next trial' button to ensure that they were prepared to advance to the next trial. Subjects were allowed to take breaks when necessary. Stimuli presentation and data collection was controlled by a REALbasic program on a Macintosh computer with a 22 inch (55.9 cm) monitor.

Procedure and stimuli

Each session consisted of a practice task and four placement tasks.[1] On practice trials, participants were presented with a number line with endpoints 0 on the left and 10 on the right. Participants were told that a number (or letter, in subsequent tasks) would appear at the top of the screen at the center of the line and it was their job to move the mouse to indicate where that number (or letter) fell on the line relative to the endpoints. During the first three practice trials, participants were given an unlimited amount of time to respond. During the last four practice trials, participants were given 3 seconds to respond before a 'Time Expired' screen appeared and the trial ended. In the first of these trials (i.e. the fourth practice trial), participants were asked to wait until the 'Time Expired' screen appeared in order to observe how much

time was allotted for each trial. Participants were then asked to complete the tasks as quickly as possible while still responding accurately. Across tasks, each stimulus value was presented twice and the median response was used in analyses.

After the practice task, participants were presented with four line placement tasks:

Non-Standard Endpoint NP (endpoints 1639 and 2897). The Non-Standard Endpoint task, where neither endpoint ended in 0, was used to explore whether the specific endpoints involved may affect responding. Subjects were asked to place the following values on the line: 1648, 1659, 1668, 1671, 1692, 1701, 1727, 1781, 1803, 1883, 1902, 1963, 1981, 2080, 2087, 2148, 2199, 2221, 2234, 2258, 2345, 2398, 2663, 2799.

Standard Endpoint, Similar Magnitude NP (endpoints 2000 and 3000). To ensure that logarithmic responding in the non-standard endpoint task was not due to logarithmic encoding of large values in this range, adults were presented with an NP task with standard endpoints (ending in 0) involving a similar numerical magnitude as the Non-Standard Endpoint NP. Subjects placed the following values: 2019, 2028, 2034, 2060, 2062, 2097, 2130, 2140, 2149, 2218, 2236, 2291, 2332, 2339, 2349, 2360, 2372, 2392, 2421, 2434, 2681, 2756, 2958, 2977.

Standard Endpoint, Similar Range NP (endpoints 0 and 1258). To explore whether responding in the Non-Standard Endpoint NP task was affected by the unorthodox range of values tested, participants were presented with an NP task with endpoints creating an identical range as the Non-Standard Endpoint task (i.e. 2897 – 1639 = 1258), but involving a standard endpoint, 0. Participants placed the following values: 25, 92, 158, 176, 180, 203, 206, 285, 310, 383, 397, 402, 455, 478, 491, 540, 547, 558, 589, 591, 795, 990, 1177, 1218.

Alphabet Placement (AP) Task (endpoints A to Z). This is discussed in Experiment 2.

All participants were presented with half of the Standard Endpoint, Similar Range (0 to 1258) trials at the beginning of the session and the other half of trials at the end of the session in order to test for fatigue or practice effects that may have impacted model fits at the end of the session. The order of the other placement tasks was counterbalanced across participants.

Data analyses

Our analyses followed the method of Siegler and colleagues (e.g. Siegler & Opfer, 2003) in order to

---

[1] Two additional placement tasks (a backwards NP line (1258–0) and a shortened alphabet line task (D–T)) were presented for the purposes of pilot data collection and will not be discussed further. Both tasks resulted in linear responding.

directly compare our findings with those frequently reported in the literature using the same methods and procedure. Thus, linear and logarithmic functions were fit to the median estimates from each task and best fit was determined by a t-test on the sum of squares of the residuals (errors) of each model. The coefficient of determination (R^2) is reported for the group data of each model. In addition, to complement the group-level statistics, non-parametric statistics at the individual level were also computed (as per Siegler & Opfer, 2003). The number of participants with a higher R^2 value (suggesting a better model fit) for the best-fit logarithmic or linear model is reported along with a binomial statistic reporting the significance of this number being different from chance (50% of subjects).

In addition to the logarithmic and linear fits, data were also fit with unbounded, 1-cycle, and 2-cycle power models in secondary analyses (using the supplemental material of Slusser *et al.*, 2013; see Appendix A for corresponding equations and parameters and Appendix B for the full set of secondary analyses). These analyses, however, did not reveal consistent strengths for one model fit over another across tasks and thus will not be discussed further in the main text. It is important to note that because the primary purpose of this study was to test claims put forth by Siegler and colleagues (e.g. Siegler & Opfer, 2003) regarding logarithmic and linear responding, we replicated their methods by testing values chosen from a logarithmic distribution (sampling more values at the lower end of the range). This methodological choice allowed for a direct comparison of our data to those of Siegler and colleagues in our primary analysis. However, Barth and colleagues (e.g. Barth *et al.*, 2011) have criticized this uneven distribution as biasing the data towards a logarithmic or linear fit and away from proper fitting via power models. We consider this criticism to be a valid one and thus caution against over-interpretation of the lack of specificity provided by these secondary analyses (see Appendix B).

Results

To ensure that the response patterns did not vary over the course of the session due to fatigue and/or practice effects, fits of the logarithmic and linear models to data from the 0 to 1258 task presented at the beginning of the session were compared to fits of data from the same task administered at the end of the session. These analyses revealed that performance between the time-points was not significantly different, as determined by comparing the residuals (errors) from the linear ($t(23) = 0.675$, $p = .506$) and logarithmic ($t(23) = 0.304$, $p = 0.764$) fits of the data from the two time-points. Thus, data from both

time-points were combined and median estimates were used in all further analyses.

Consistent with previous investigations of adult NP task performance (e.g. Siegler *et al.*, 2009), results of all tasks involving *standard* endpoints resulted in linear responding. That is, responding in both the 0–1258 NP task ($R^2_{lin} = 0.995$, $R^2_{log} = 0.732$, $t(23) = 5.141$, $p < .001$, 38/40 linear, *binomial p < .001*) and the 2000–3000 *NP task* ($R^2_{lin} = 0.981$, $R^2_{log} = 0.967$, $t(23) = 3.852$, $p < .005$, 38/40 linear, *binomial p < .001*) were significantly better fit by a linear model. In contrast, data from the *Non-Standard Endpoint* NP task (1639 to 2897) were better fit by a logarithmic model ($R^2_{log} = 0.988$, $R^2_{lin} = 0.982$, $t(23) = 2.691$, $p < .015$, 23/40 logarithmic, *binomial p > .2*).

Experiment 1b

Results of Experiment 1a suggest that logarithmic responding results from the increased difficulty in mentally manipulating non-standard endpoints (1639 to 2897). Because these are the first findings of their kind, in Experiment 1b, we wanted to verify the robustness of this finding by running the same task with a different group of adult subjects. At the same time, we corrected for a minor discrepancy across conditions in Experiment 1a. Post hoc it was realized that the distribution of stimuli used in the Non-Standard Endpoint condition differed slightly from the distribution of stimuli in the control tasks (2000 to 3000 and 0 to 1258) in Experiment 1a. Whereas the distribution of stimuli in the Non-Standard Endpoint condition mimicked that of previous studies (e.g. Siegler & Opfer, 2003), the distributions in our control conditions included very few values in the first quarter of the range. Given the possibility that the distribution of stimuli chosen in the different conditions affected the resulting behavior (Duffy, Huttenlocher & Crawford, 2006), in Experiment 1b, the tasks were repeated with an identical distribution of stimuli across tasks. In doing so, Experiment 1b also served as a replication of Experiment 1a.

Method

Participants

Participants were 25 adults recruited from the Boston, MA area, tested either at the Boston Children's Museum or at Boston College. Five adults were excluded from analysis for not completing the task ($N = 4$) and for interference during the task ($N = 1$). Thus, 20 adults (nine males; $M = 24.9$ years; Range: 18 to 41 years) were included in the final analysis.

Experimental design and materials

The design was identical to that of Experiment 1a except that, because the task was not always conducted in the lab, participants were offered noise-cancelling headphones which played white noise to reduce the likelihood of external distractions during the task.

Procedure and stimuli

The procedure was identical to that of Experiment 1a; only the distribution of stimulus values in the control conditions was changed to ensure consistency across conditions.

Non-Standard Endpoint NP (1639 to 2897). This was identical to Experiment 1a.

Standard Endpoint, Similar Magnitude NP (2000 to 3000). Each of the following values was presented twice: 2007, 2016, 2023, 2025, 2042, 2049, 2070, 2113, 2131, 2195, 2210, 2259, 2273, 2352, 2358, 2407, 2448, 2465, 2476, 2495, 2564, 2607, 2819, 2928.

Standard Endpoint, Similar Range NP (0 to 1258). Each of these values was presented once at the beginning of the experimental session and again once at the end of the session: 9, 20, 29, 32, 53, 62, 88, 142, 164, 244, 263, 324, 342, 441, 448, 509, 560, 582, 595, 619, 706, 759, 1024, 1160.

Alphabet Placement (AP) Task (endpoints A to Z). This task was repeated in Experiment 1b to better match the stimuli values used for the children in Experiment 2. All AP tasks will be discussed in Experiment 2.

All participants were presented with the Standard Endpoint, Similar Range NP task once at the beginning and once at the end of the session. The order of the other NP tasks was counterbalanced across participants and the AP task was always presented second to last.

Data analyses

All analyses were identical to Experiment 1a.

Results

As in Experiment 1a, there was no difference in performance patterns between the 0 to 1258 tasks at the beginning and end of the experiment (linear fit: $t(23) = 0.048$, $p = .962$; logarithmic fit: $t(23) = 0.807$, $p = .428$). Thus, data from both time-points were combined and median estimates were used for further analyses.

Results of Experiment 1b replicated those of Experiment 1a. In particular, responses in both the 0–1258 NP task ($R^2_{lin} = 0.987$, $R^2_{log} = 0.727$, $t(23) = 5.470$, $p < .001$, 20/20 linear, *binomial p* < .001) and the 2000–3000 NP task ($R^2_{lin} = 0.985$, $R^2_{log} = 0.971$, $t(23) = 4.636$, $p < .001$, 20/20 linear, *binomial p* < .001) were significantly better fit by a linear model (Figure 1). Importantly, data from the Non-Standard Endpoint task were again better fit by a logarithmic model ($R^2_{log} = 0.981$, $R^2_{lin} = 0.970$, $t(23) = 2.365$, $p < .03$, 14/20 logarithmic, *binomial p* < .06). Thus results indicate that differences in estimation patterns observed in Experiment 1a were not due to differences in the distribution of values chosen, but rather to differences in the specific endpoint values.
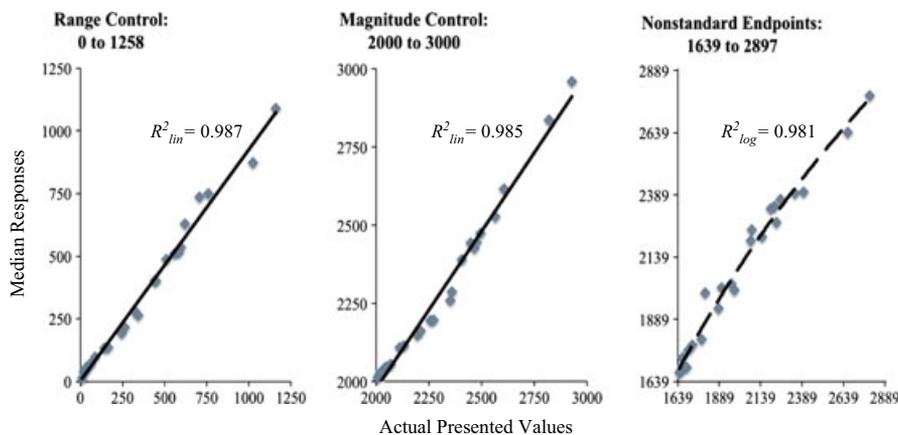

Figure 1 *Best fitting function (linear or logarithmic) for data from adults in Experiment 1b. Data from the Range Control and Magnitude Control conditions were best fit by a linear function, while the Non-Standard Endpoint condition was best fit by a logarithmic function.*

Discussion

Results of Experiments 1a and 1b suggest that adults respond logarithmically when presented with non-standard endpoints that are less familiar and possibly more difficult to mentally manipulate. Adults in both experiments responded logarithmically when presented with non-standard endpoints (1639 to 2897), despite responding linearly on tasks involving standard endpoints with the same numerical range (0 to 1258) and a similar magnitude (2000 to 3000). Importantly, these findings support our hypothesis that behavior on the NP task does not directly reflect our internal representation of number, as response patterns were entirely independent of the numerical magnitudes involved. Rather, data suggest that logarithmic versus linear patterns of responding result from the familiarity of the endpoints and ease with which mental calculations involving the endpoint values can be performed.

Relatedly, our data may also account for the context effects reported by previous studies (e.g. Siegler & Opfer 2003; Siegler et al., 2009), in which, for example, children respond linearly on a 0 to 10 line but logarithmically on a 0 to 100 line. That is, when endpoints on the NP task are less familiar (e.g. 100 for kindergarteners and 1000 for 2nd graders), making it difficult for children to correctly compare the relative distances between the target value and the two endpoints, logarithmic performance results. In contrast, when endpoints are more familiar (e.g. 100 for 2nd graders 1000 for 4th graders; Seigler et al., 2009), thereby facilitating mental computations in order to appropriately judge the relative distances, children respond linearly. Our data parallel this pattern in adults, where they respond logarithmically on tasks with unfamiliar endpoints and linearly on the magnitude and range control tasks with familiar endpoints.

In sum, we contend that, contrary to claims made by Siegler, Opfer and colleagues (e.g. Siegler & Opfer, 2003), behavior on the NP task does not directly reflect the underlying representation of number as it is unaffected by the range or magnitude of values tested. Instead, it is at least partially indicative of fluency in working with the particular endpoint values presented.

Experiment 2

In our second experiment, we provide a direct test of our second hypothesis that NP task performance is dependent upon fluency with the specific ordering of values within the range. To do so, we explored the development of ordering the alphabet, a non-numerical sequence without an inherent underlying magnitude. Whereas linear spacing is mathematically inherent to the numerical sequence such that successive numeric values are spaced equally, (i.e. $8 - 7 = 2 - 1$), there is no property of the alphabet that requires alphabetical ordering to be linear. Although the alphabet has a standard ordering, the sequence itself does not imply that, for example, the distance between A and B is the same as the distance between E and F. There is no 'distance' between A and B when considering the alphabetical sequence as simply a string of ordered characters.

In Experiment 2, we analyzed children's and adults' behavior on an alphabet placement (AP) task requiring participants to place letters of the alphabet on a line with endpoints A and Z. For comparison, we analyzed children's behavior on a standard NP task. If a similar developmental trajectory is found in placement behavior for a sequence with inherent mathematically defined magnitude (e.g. $4 - 1 = 3$, $8/2 = 4$, etc.) and a sequence with no apparent magnitude whatsoever (e.g. M/F = ?, B + E = ?, etc.), then this will provide strong evidence against claims that the logarithmic to linear shift in responding reflects a developmental shift in the underlying representation of magnitude.

Methods

Participants

One hundred and twelve children between the ages of 6 and 11 years were recruited from the Boston, MA area and tested in our laboratory on the main campus of Boston College, at the Museum of Science (Boston, MA), and at local schools. Participants were excluded from data analysis for one or both of the tasks based on the following criteria: (i) interference from a parent or another child ($N = 2$ excluded from both tasks); (ii) computer error ($N = 1$ excluded from both tasks and an additional $N = 3$ excluded from just the NP task); (iii) failure to complete the task ($N = 1$ excluded from both tasks); (iv) excessive number of trials in which responding exceeded the allotted response time[2] ($N = 2$ excluded from the NP task and a different $N = 3$ excluded from the AP task); and (v) behavior classified as not understanding or following instructions, evidenced by an R^2 value of less than 0.2 for both the linear and logarithmic fits of the data (indicative of non-systematic responding; $N = 4$ excluded from both tasks, with an

---

[2] Because the NP task involved a greater number of trials than the AP task, we applied a different criterion for exclusion for the two tasks. Participants with three time-outs or more on the AP task were excluded from that task and those with four time-outs or more on the NP task were excluded from data analyses for that task.

additional $N = 4$ excluded from the NP task only and $N = 2$ excluded from the AP task only). In total, 13 children were excluded from AP task analyses and 17 were excluded from NP task analyses.

A total of 95 participants were included in NP task analyses, grouped by age: 6–7 years ($n = 33$; $M = 6$y11m; Range: 5y10m to 7y10m; 19 males), 8–9 years ($n = 44$; $M = 9$y3m; Range: 8y1m to 9y10m; 25 males), and 10–11 years ($n = 18$; $M = 10$y8m; Range: 10y0m to 11y5m; six males).

For the AP task, a total of 99 participants were included in analyses, grouped by age: 6–7 years ($n = 32$; $M = 7$y0m; Range: 5y10m to 7y10m; 16 males), 8–9 years ($n = 46$; $M = 9$y2m; Range: 8y1m to 9y10m; 27 males), and 10–11 years ($n = 21$; $M = 10$y8m; Range: 10y0m to 11y5m; six males). In addition, data from 40 adults in Experiment 1a and 20 adults in Experiment 1b who participated in the AP task will be discussed here along with child data.

Experimental design and procedure

Participants were given a practice task (identical to Experiment 1), an NP task (0 to 1000) and an AP task (A to Z) in a single experimental session (order counterbalanced). The remainder of the set-up, experimental design, and procedure was identical to that of Experiment 1, except that child participants were given up to 15 seconds to respond on every trial (instead of the 3 second limit used for adults).

NP task (0 to 1000). Typical behavior on this task and for this range has been well documented (e.g. Siegler *et al.*, 2009). The purpose of this task was to replicate previous research with this group of participants. The following values were presented once each in a random order: 19, 25, 67, 83, 92, 158, 176, 180, 203, 206, 285, 310, 346, 383, 397, 402, 432, 455, 478, 491, 624, 795, 812, 990.

AP task (A to Z). This task was presented in all three experiments. Adult participants in Experiment 1a were presented with the following values twice in a random order: B, C, D, E, F, G, H, I J, K, L, M, P, R, and Y. Adults in Experiment 1b and children in Experiment 2 were presented each of the following values once in a random order: C, D, E, F, G, H, I, J, K, N, S.

Data analyses

For purposes of data coding for the AP task, 'A' was coded as '0' and 'Z' was coded as '25' to mimic previous NP data analyses. As in Experiment 1, all group analyses were performed on median responses and all analyses and reported statistics are identical to Experiment 1.

Results

Unsurprisingly, results of the NP task replicated previous research (e.g. Siegler & Booth, 2004; Siegler & Opfer, 2003; Siegler *et al.*, 2009), revealing a shift from logarithmic to linear response distributions with age. Data from the 6–7-year age group were significantly better fit by the logarithmic model ($R^2_{log} = 0.941$, $R^2_{lin} = 0.824$, $t(23) = 2.699$, $p = .013$, 25/33 logarithmic, *binomial* $p < .01$). For older children, however, the pattern was reversed, with data from the 8–9-year age group ($R^2_{lin} = 0.990$, $R^2_{log} = 0.726$, $t(23) = 5.331$, $p < .001$; 38/44 linear, *binomial* $p < .001$) and the 10–11-year age group ($R^2_{lin} = 0.992$, $R^2_{log} = 0.725$, $t(23) = 5.512$, $p < .00005$; 17/18 linear, *binomial* $p < .001$) being significantly better fit by the linear model.

In the AP task, a remarkably similar developmental trend was found showing a progression from logarithmic behavior to linear behavior with age. Data from the 6–7-year age group was significantly better fit with the logarithmic model ($R^2_{log} = 0.981$, $R^2_{lin} = 0.925$, $t(10) = 2.319$, $p = .043$; 19/32 logarithmic, *binomial* $p = .189$). In the 8–9-year group, mixed behavior led to insignificant differences between the linear and logarithmic models ($R^2_{lin} = 0.973$, $R^2_{log} = 0.951$, $t(10) = 0.322$, $p = .754$; 19/46 logarithmic, *binomial* $p = .908$). However, data from the 10–11-year age group ($R^2_{lin} = 0.992$, $R^2_{log} = 0.928$, $t(10) = 2.603$, $p = .026$; 14/21 linear, *binomial* $p = .0946$) and both adult groups (Experiment 1a: $R^2_{lin} = 0.996$, $R^2_{log} = 0.824$, $t(14) = 4.560$, $p < .001$; 40/40 linear; Experiment 1b: $R^2_{lin} = 0.996$, $R^2_{log} = 0.892$, $t(10) = 3.657$, $p < .005$; 17/20 linear) were significantly better fit with the linear model (see Figure 2).

Discussion

Data from both the number placement and alphabet placement tasks show a gradual shift from logarithmic to linear responding between 6 and 11 years of age, with linear placement of stimuli for the AP task in adulthood. NP task results replicate patterns previously reported in the literature (Siegler *et al.*, 2009). Data from the AP task extend those recently reported by Berteletti, Lucangeli and Zorzi (2012), who reported the beginning of a similar developmental shift in non-numerical ordering in young children. However, Berteletti *et al.* (2012) only provided evidence for logarithmic behavior in young children (around 6 to 7 years) and mixed behavior in slightly older children (around 8 to 9 years) while the present data reveal that consistent linear behavior in an AP task is attained within older children (10 to 11 years) and is maintained into adulthood. Indeed, the evidence indicates that the full developmental trajectory of
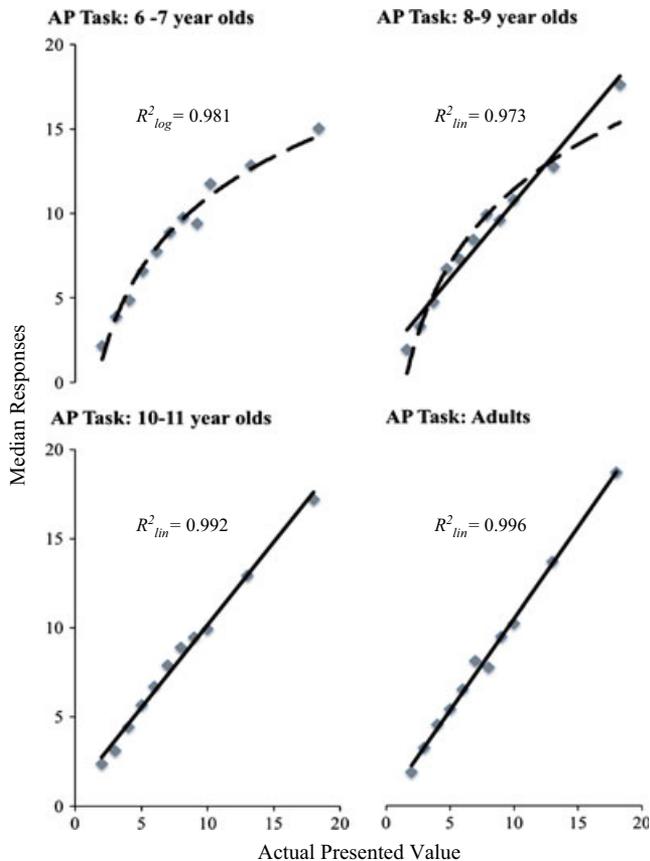
AP Task: 6 -7 year olds

$R^2_{log}= 0.981$

AP Task: 8-9 year olds

$R^2_{lin}= 0.973$

AP Task: 10-11 year olds

$R^2_{lin}= 0.992$

AP Task: Adults

$R^2_{lin}= 0.996$

Median Responses

Actual Presented Value

Figure 2 *Best fitting function(s) (linear or logarithmic) for data from kids and adults in Alphabet Placement Tasks. Six- to 7-year-olds were best fit by a logarithmic function, 8- to 9-year-olds were equally well fit by both linear and logarithmic functions, and 10- to 11-year-olds and adults were better fit by a linear function.*

alphabet spatial estimation involves a logarithmic-to-linear shift in behavior between 6 and 11 years of age.

Although it has been suggested that the transition from logarithmic to linear behavior reflects a shift in the underlying representation of number (Siegler & Opfer, 2003), data from our alphabet task again call this theory into question. The alphabet is a list of all the letters used in the English language, with an arbitrary order. The letters in the alphabet have no inherent numerical (or otherwise quantitative) magnitude, and are not inherently linearly spaced, nor are they subject to the laws of arithmetic. Despite this, our data reveal an identical, systematic shift from logarithmic spacing to linear spacing with age for letters and for numbers, calling into question whether the pattern seen in the number line task is truly indicative of a change in the underlying representation of number. Instead, we contend that the developmental shift found in both cases is indicative of the acquisition of a more general tendency to linearly

space ordered lists, once the ordering of the entire sequence is fully mastered. Although our youngest subjects (6–7 years) were clearly familiar with the alphabet (as evidenced by non-random responding in the AP task), they likely had a less fluid or automatic understanding of alphabetical order than older children and adults. While most adults simply *know* that 'S' comes before 'T', for example, younger children may need to sing through the alphabet song in order to determine the relative ordering of these values; that is, their understanding of the ordering is less automatic and more deliberate. This lack of fluency with the ordering of the values results in a greater focus on values early in the sequence, which were likely learned first and practiced more. This shift in attention decompresses the items within focus and compresses the items out of focus (Lourenco & Longo, 2009), leading to the logarithmic pattern of responses. With age comes greater experience with the sequence, which leads to greater automaticity in processing the ordering of the sequence and a more distributed focus across the entire range, resulting in a linear pattern of responding. In sum, a logarithmic-to-linear developmental trend in responding may simply reflect automaticity with the ordering of the sequence, with a general tendency towards linearizing familiar, ordered lists.

General discussion

The number line placement task has been used extensively in research to investigate numerical estimation ability in children (e.g. Booth & Siegler, 2008; Berteletti, Lucangeli, Piazza, Dehaene & Zorzi, 2010) and to investigate the underlying representation of number (e.g. Booth & Siegler, 2006; Opfer & DeVries, 2008; Siegler & Opfer, 2003). Linear performance on the NP task has been found to be predictive of more advanced math ability (Siegler & Booth, 2004), arithmetic ability (Booth & Siegler, 2008), and memory for numbers (Thompson & Siegler, 2010), suggesting that performance on the NP task is indicative of an underlying competence with numbers and/or mathematical competence. Siegler and colleagues suggest that the predictive ability of the NP task stems from the insight gained into the structure of the underlying representation of number. In contrast, evidence from the present study suggests that NP task performance does not provide insight into the structure of the underlying representation of number. Rather, the task may be related to mathematical competence as it reflects a level of mastery with particular manipulations of the endpoints and fluency with the (not necessarily mathematical) ordering of the values in the interval.

Given evidence suggesting children's representations of number become increasingly precise across development (Lipton & Spelke, 2004; Halberda & Feigenson, 2008), we acknowledge that it is possible that there is a shift in their underlying representation of number during early childhood. Our data do not rule out the possibility that children undergo a developmental shift in the underlying representation of number. However, our results do make clear that if such a developmental shift exists, it is *not* reflected in performance on the NP task. In addition, our data do not call into question the predictive ability of the NP task, as it could be measuring other important aspects of numerical competence, such as fluency with the ordering of the values and/or an ability to implicitly make proportion computations across the range of values presented.

Previous studies have shown that preventing the use of proportional information may alter responding on the NP task. Cohen and Blanc-Goldhammer (2011) suggest that an unbounded number line task may provide a more pure mapping from internal representation of number to external behavior as it prohibits proportion-based responding. Furthermore, Rips (2013) disrupted linear behavior in adults by using endpoints of a zillion and a google, which are extremely unfamiliar and possibly even unknown to the participant. Our data further extend these results by providing evidence of an even simpler disruption of linear strategy when encountering values that may be somewhat less familiar and not necessarily easy to mentally manipulate (i.e. 1639 and 2897).

If it is the case that responding is dictated by proportion judgments and is *not* indicative of the format of encoding of subjective number, then why is it that responding in some circumstances appears *logarithmic*? Karolis, Iuculano and Butterworth (2011) have suggested that logarithmic responding may result from an individual's tendency to overestimate in the lower end of the range and underestimate values in the upper end of the range (termed central tendency). This effect may be related to an attentional shift, such that the portion of the number line that is in focus is decompressed, while the out-of-focus portion is compressed (Lourenco & Longo, 2009). Thus, in situations of uncertainty (e.g. in a less familiar context) an attentional focus on the first half of the number line (because of familiarity and reading order) and our more general central tendency may result in logarithmic responding (e.g. Karolis *et al.*, 2011; Lourenco & Longo, 2009). Similarly, Petitto (1990) provided evidence of a sequence-to-proportion shift in estimation strategy, which suggests that children's early estimation strategies only use one endpoint and ignore distance, simply placing values sequentially, with more advanced strategies incorporating an understanding of

distance and equal spacing between values. In addition, attentional resources have been implicated in non-symbolic number line estimation such that adults participating in dual-task (cognitively taxing) paradigms are more likely to space values logarithmically (Anobile, Cicchini & Burr, 2012). Consistent with these ideas, it may be the case that when individuals are presented with less familiar endpoints and/or a less familiar ordered sequence, determining where values are placed along the line becomes a cognitively demanding task, resulting in data patterns consistent with central tendency.

In conclusion, our findings cast doubt upon claims that the NP task allows for direct interpretation of the underlying representation of number. Instead, results suggest that responding is influenced by the use of proportional information, which is dependent both upon an ability to mentally manipulate the endpoints and upon fluency (i.e. automaticity) with the relative ordering of the sequence of values within the range. Data from two experiments support these claims by showing that (1) logarithmic behavior can be induced in adults by altering the endpoints of the interval, thereby disrupting their use of proportion-based strategies and (2) the development of non-numerical sequence ordering in children shows a similar logarithmic to linear shift. Thus, the logarithmic-to-linear developmental shift in behavior reported in the number line estimation task appears to reflect the use of strategies employed across a familiar range of values, rather than a deep, conceptual change in the mental representation of number.

The task of clarifying the format of the mapping between objective and subjective number is a formidable one, but one that is worth pursuing. This information will not only shed light on how number is represented in the mind, but has the potential to inform both educational practices and neuroscientific investigations. Results of the current study call into question evidence that has been put forth to support claims of a shift in the underlying representation of number across childhood. While data do not rule out the possibility of a developmental shift, our findings reveal that NP task data are unable to speak to claims regarding the underlying representation. Instead, our study points to the need for further research in this domain to determine exactly how it is that number is represented in the mind across development.

Acknowledgements

References

Anobile, G., Cicchini, G.M., & Burr, D.C. (2012). Linear mapping of numbers onto space requires attention. *Cognition*, **122**, 454–459. doi: 10.1016/j.cognition.2011.11.006

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

Barth, H., Slusser, E., Cohen, D., & Paladino, A. (2011). A sense of proportion: commentary on Opfer, Siegler and Young. *Developmental Science*, **14**, 1205–1206. doi: 10.1111/j.1467-7687.2011.01081.x

Berteletti, I., Lucangeli, I., Piazza, M., Dehaene, S., & Zorzi, M. (2010). Numerical estimation in preschoolers. *Developmental Psychology*, **46**(2), 545–551. doi: 10.1037/a0017887

Berteletti, I., Lucangeli, D., & Zorzi, M. (2012). Representation of numerical and non-numerical order in children. *Cognition*, **124**, 304–313.

Booth, J., & Siegler, R. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology*, **41**(6), 189–201. doi: 10.1037/0012-1649.41.6.189

Booth, J., & Siegler, R. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development*, **79**(4), 1016–1031. doi: 10.1111/j.1467-8624.2008.01173.x

Brannon, E., Wusthoff, C., Gallistel, C., & Gibbon, J. (2001). Numerical subtraction in the pigeon: evidence for a linear subjective number scale. *Psychological Science*, **12**(3), 238–243. doi: 10.1111/1467- 9280.00342

Buckley, P., & Gillman, C. (1974). Comparisons of digits and dot patterns. *Journal of Experimental Psychology*, **103**(6), 1131–1136.

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

Cantlon, J., Cordes, S., Libertus, M., & Brannon, E. (2009). Comment on 'Log or linear? Distinct intuitions of the number scale in western and Amazonian indigene cultures'. *Science*, **323**, 38. doi: 10.1126/science.1164773

Cohen, D., & Blanc-Goldhammer, D. (2011). Numerical bias in bounded and unbounded number line tasks. *Psychonomic Bulletin & Review*, **18**, 331–338. doi: 10.3758/s13423-011-0059-z

Cordes, S., Gallistel, C.R., Gelman, R., & Latham, P. (2007a). Nonverbal arithmetic in humans: light from noise. *Perception & Psychophysics*, **69**(7), 1185–1203.

Cordes, S., King, A.P., & Gallistel, C.R. (2007b). Time left in the mouse. *Behavioural Processes*, **74**, 142–151.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, **22**, 1–42.

Dehaene, S. (2001). Subtracting pigeons: logarithmic or linear? *Psychological Science*, **12**(3), 244–246.

Dehaene, S. (2003). The neural basis of the Weber-Fechner law: a logarithmic mental number line. *Trends in Cognitive Sciences*, **7**(4), 145–147.

Dehaene, S., Izard, V., Spelke, E., & Pica, P. (2008). Log or linear? Distinct intuitions of the number scale in western and Amazonian indigene cultures. *Science*, **320**, 1217–1220. doi: 10.1126/science.1156540

Duffy, S., Huttenlocher, J., & Crawford, L.E. (2006). Children use categories to maximize accuracy in estimation. *Developmental Science*, **9**(6), 597–603.

Gallistel, C.R., Brannon, E.M., Gibbon, J., & Wusthoff, C.J. (2001). Response to Dehaene. *Psychological Science*, **12**(3), 247.

Geary, D.C., Bailey, D.H., & Hoard, M.K. (2009). Predicting mathematics achievement and mathematical learning disability with a simple screening tool: the number sets task. *Journal of Psychoeducational Assessment*, **27**, 295–279. doi: 10.1177/0734282908330592

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, **84**(3), 279–325.

Gibbon, J., & Church, R. (1981). Time left: linear versus logarithmic subjective time. *Journal of Experimental Psychology: Animal Behavior Processes*, **7**(2), 87–108.

Gibbon, J., & Church, R. (1990). Representation of time. *Cognition*, **37**, 23–54.

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the 'number sense': the approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, **44**(5), 1457–1465. doi: 10.1037/a0012682

Halberda, J., Mazzocco, M.M.M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, **455**(2), 665–668. doi: 10.1038/nature07246

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

Jordan, K., & Brannon, E. (2006). A common representational system governed by Weber's law: nonverbal numerical similarity judgments in 6-year-olds and rhesus macaques. *Journal of Experimental Child Psychology*, **95**, 215–229. doi: 10.1016/j.jecp.2006.05.004

Karolis, V., Iuculano, T., & Butterworth, B. (2011). Mapping numerical magnitudes along the right lines: differentiating between scale and bias. *Journal of Experimental Psychology: General*, **140**(4), 693–706. doi: 10.1037/a0024255

Laski, E., & Siegler, R. (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. doi: 10.1111/j.1467-8624.2007.01087.x

Lipton, J.S., & Spelke, E.S. (2004). Discrimination of large and small numerosities by human infants. *Infancy*, **5**(3), 271–290.

Lourenco, S.F., & Longo, M.R. (2009). Multiple spatial representations of number: evidence for co-existing compressive and

linear scales. *Experimental Brain Research*, **193**, 151–156. doi: 10.1007/s00221-008-1698-9

Moyer, R., & Landauer, T. (1967). Time required for judgements of numerical inequality. *Nature*, **215**, 1519–1520.

Nieder, A., & Miller, E. (2003). Coding of cognition magnitude: compressed scaling of numerical information in the primate prefrontal cortex. *Neuron*, **37**, 149–157.

Opfer, J., & DeVries, J. (2008). Representational change and magnitude estimation: why young children can make more accurate salary comparisons than adults. *Cognition*, **108**, 843–849. doi: 10.1016/j.cognition.2008.05.003

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

Opfer, J., & Thompson, C. (2008). The trouble with transfer: insights from microgenetic changes in the representation of numerical magnitude. *Child Development*, **79**(3), 788–804. doi: 10.1111/j.1467- 8624.2008.01158.x

Petitto, A. (1990). Development of numberline and measurement concepts. *Cognition and Instruction*, **7**(1), 55–78.

Rips, L. (2013). How many is a zillion? Sources of number distortion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **39**(4), 1257–1264. doi: 10.1037/a0031143

Schneider, M., Grabner, R.H., & Paetsch, J. (2009). Mental number line, number line estimation, and mathematical achievement: their interrelations in grades 5 and 6. *Journal of Educational Psychology*, **101**(2), 359–372. doi: 10.1037/a0013840

Siegler, R., & Booth, J. (2004). Development of numerical estimation in young children. *Child Development*, **75**(2), 428–444. doi: 10.1111/j.1467-8624.2004.00684.x

Siegler, R., & Booth, J. (2005). Development of numerical estimation: a review. In J.I.D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 197–212). New York: Psychology Press.

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

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

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

Siegler, R., Thompson, C., & Opfer, J. (2009). The logarithmic-to-linear shift: one learning sequence, many tasks, many time scales. *Mind, Brain, and Education*, **3**(3), 143–150. doi: 10.1111/j.1751- 228X.2009.01064.x

Siegler, R.S., Thompson, C.A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology*, **62**, 273–296. doi: 10.1016/j.cogpsych.2011.03.001

Slusser, E., Santiago, R., & Barth, H. (2013). Developmental change in numerical estimation. *Journal of Experimental Psychology: General*, **142**(1), 193–208. doi: 10.1037/a0028 560

Sullivan, J., Juhasz, B., Slattery, T., & Barth, H. (2011). Adults' number-line estimation strategies: evidence from eye movements. *Psychonomic Bulletin & Review*, **18**, 557–563. doi: 10.3758/s13423-011-0081-1

Thompson, C., & Siegler, R. (2010). Linear numerical-magnitude representations aid children's memory for numbers. *Psychological Science*, **21**(9), 1274–1281. doi: 10.1177/0956797610378309

Whalen, J., Gallistel, C., & Gelman, R. (1999). Nonverbal counting in humans: the psychophysics of number representation. *Psychological Science*, **10**(2), 130–137.

Appendix A

Model fitting equations for secondary analysis (analyzed using Slusser, Santiago & Barth, 2013, supplemental Excel file)

Unbounded power model

$$y = A \times x^\beta$$

Free Parameters (2): A; β

1-Cycle power model

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

Free Parameters (1): β
Constant (1): R (fixed at range)

2-Cycle power model

$$y = R \times \left(\frac{0.5(x^\beta)}{x^\beta + (\frac{R}{2}-x)^\beta} \right) \qquad for\ x < \frac{R}{2}$$

Free Parameters (1): β
Constant (1): R (fixed at range)

$$y = R \times \left(\left(\frac{0.5(x-\frac{R}{2})^\beta}{(x-\frac{R}{2})^\beta + (R-x)^\beta} \right) + 0.5 \right) \qquad for\ x \geq \frac{R}{2}$$

Free Parameters (1): β
Constant (1): R (fixed at range)

AICc: Akaike Information Criterion (Corrected for small sample sizes)

$$AICc = AIC + \left(\frac{2K(K+1)}{N-K-1} \right)$$

K is the number of parameters in the model
N is the number of stimuli presented (i.e. the number of x values)

Appendix B

Statistical details of secondary analysis

Experiment 1a	0 to 1258 task		2000 to 3000 task		1639 to 2897 task	
Model	AICc	ΔAICc	AICc	ΔAICc	AICc	ΔAICc
Logarithmic	246.19	94.3	190.6	13.1	*172.6*	*
Linear	*151.9*	*	*177.5*	*	183.1	10.5
Unbounded Power	330.3	178.4	323.4	145.9	329.1	156.5
1-Cycle Power	187.6	35.7	185.01	7.51	191.6	19
2-Cycle Power	186.5	34.6	193.49	15.99	211.6	39

Experiment 1b	0 to 1258 task		2000 to 3000 task		1639 to 2897 task	
Model	AICc	ΔAICc	AICc	ΔAICc	AICc	ΔAICc
Logarithmic	246.3	74	186.2	26.3	*185.4*	*
Linear	*172.3*	*	170.4	10.5	196.1	10.7
Unbounded Power	333.6	161.3	323.0	163.1	329.4	144
1-Cycle Power	188.4	16.1	*159.9*	*	201.9	16.5
2-Cycle Power	263.5	91.2	176.6	16.7	215.4	30

Experiment 2: Alphabet	6 to 7 years		8 to 9 years		10 to 11 years	
Model	AICc	ΔAICc	AICc	ΔAICc	AICc	ΔAICc
Logarithmic	*−8.8*	*	−2.5	5.18	7.9	33
Linear	6.1	14.9	4.07	11.75	−16.6	8.5
Unbounded Power	−1.11	7.7	*−7.62*	0.06	*−23.7*	1.4
1-Cycle Power	−0.87	7.93	−7.68	*	−25.1	*
2-Cycle Power	11.47	20.3	3.68	11.36	−9.1	16

Alphabet Adults	Adults (Exp 1a)		Adults (Exp 1b)	
Model	AICc	ΔAICc	AICc	ΔAICc
Logarithmic	33.9	55.5	34.4	54.9
Linear	−21.0	0.6	−20.5	*
Unbounded Power	−21.6	*	−18.6	1.9
1-Cycle Power	−13.89	7.71	−12.7	7.8
2-Cycle Power	−13.92	7.68	−10.6	9.9

Note: Statistically preferred model (as determined by the lowest AICc) indicated by asterisk (*). All other models are tested for support by being compared to the preferred model. For interpretation, the following guidelines of Burnham and Anderson (2004, p. 271) may be used: 'models having $\Delta \leq 2$ have substantial support (evidence), those in which $4 \leq \Delta \leq 7$ have considerably less support, and models having $\Delta > 10$ have essentially no support'.

Mathematically, the proportion judgment model allows for an increase in accuracy with development and education by adjusting the parameters of the power function. An unbounded power function is consistent with behavior that does not include an understanding of the large (right) endpoint. On the other hand, bounded power models assume an understanding of the endpoints and can further be specified to include additional anchor points. For example, the 1-cycle power model is consistent with the use of two anchor points (namely, the endpoints) while the 2-cycle power model is consistent with the use of three anchor points (the endpoints and the midpoint; Holland & Dyre, 2000; Slusser *et al.*, 2013). In our analyses, data do not provide clear evidence for or against the proportion-judgment account, overall. However, we believe this may be due to our sampling from a logarithmic distribution of test values (as modeled from Siegler & Opfer, 2003), a methodology that has been criticized by Barth and colleagues (e.g. Barth *et al.*, 2011) as undermining the strength of the proportion-judgment account.