Cognitive Psychology 107 (2018) 44–66

Contents lists available at ScienceDirect

Cognitive Psychology

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

Dynamics and development in number-to-space mapping

Dan Kim, John E. Opfer

The Ohio State University, 255 Psychology Building, Columbus, OH 43210, USA

A R T I C L E I N F O

Keywords:
Cognitive development
Number-line estimation
Number-to-space mapping
Logarithmic compression

A B S T R A C T

Young children's estimates of numerical magnitude increase approximately logarithmically with actual magnitude. The conventional interpretation of this finding is that children's estimates reflect an innate logarithmic encoding of number. A recent set of findings, however, suggests that logarithmic number-line estimates emerge via a dynamic encoding mechanism that is sensitive to previously encountered stimuli. Here we examine trial-to-trial changes in logarithmicity of numerosity estimates to test an alternative dynamic model (D-MLLM) with both a strong logarithmic component and a weak response to previous stimuli. In support of D-MLLM, first-trial numerosity estimates in both adults (Study 1, 2, 3, and 4) and children (Study 4) were strongly logarithmic, despite zero previous stimuli. Additionally, although numerosity of a previous trial affected adults' estimates, the influence of previous numbers always accompanied the logarithmic-to-linear shift predicted by D-MLLM. We conclude that a dynamic encoding mechanism is not necessary for compressive mapping, but sequential effects on response scaling are a possible source of linearity in adults' numerosity estimation.

1. Introduction

Mapping number to space is fundamental for everyday life. Humans exhibit spontaneous associations between number and space even in early infancy (de Hevia, Izard, Coubart, Spelke, & Streri, 2014; de Hevia & Spelke, 2010, 2013; Lourenco & Longo, 2010), as do other species, such as chimpanzees and newborn chicks (Adachi, 2014; Drucker & Brannon, 2014; Rugani, Vallortigara, Priftis, & Regolin, 2015; see McCrink & Opfer, 2014, for review). The nature of this mapping is not limited to non-symbolic numbers (i.e., numerosity), extends to symbolic numbers, and improves dramatically with age and education (Ashcraft & Moore, 2012; Dehaene, Izard, Spelke, & Pica, 2008; Geary, Hoard, Nugent, & Byrd-Craven, 2008; Gunderson, Ramirez, Beilock, & Levine, 2012; Hurst, Leigh Monahan, Heller, & Cordes, 2014; McCrink & Opfer, 2014; Opfer & Furlong, 2011; Siegler & Booth, 2004; Siegler & Opfer, 2003). However, it is still controversial whether approximately logarithmic compression in number-to-space mappings should be attributed to primitive, logarithmic encoding of numerical information (Dehaene, 2003, 2007; Nieder & Merten, 2007; Siegler & Opfer, 2003) or to a dynamic encoding mechanism that is sensitive to the statistics of previously encountered numbers (Anobile, Cicchini, & Burr, 2012; Cicchini, Anobile, & Burr, 2014).

1.1. Evidence for compressive number-to-space mapping

Compressive number-to-space mappings are perhaps most evident in developmental studies of number-line estimation, such as when children estimate the location of a number (e.g., 15) on a line flanked by two other numbers (e.g., 0 and 30) (Fig. 1A). In early development, estimates of numbers increase as a logarithmic function of the number to be estimated (for review, see Kim & Opfer,

E-mail addresses: kim.3839@osu.edu (D. Kim), opfer.7@osu.edu (J.E. Opfer).

https://doi.org/10.1016/j.cogpsych.2018.10.001
Accepted 11 October 2018
Available online 12 November 2018

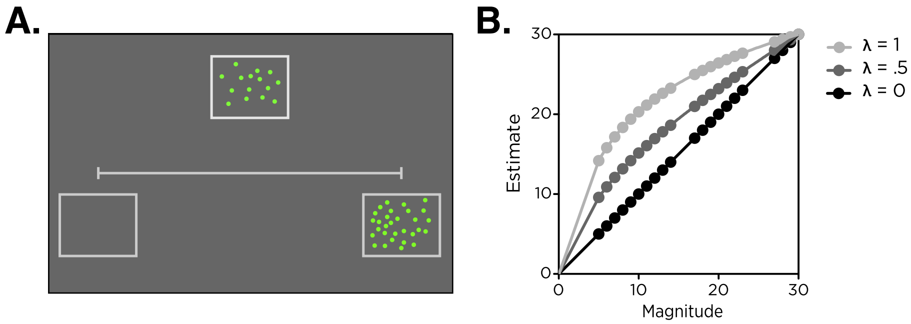

Fig. 1. Illustration of a non-symbolic number-line task (A) and responses predicted by the MLLM with various degrees of logarithmicity (B).

2017). With schooling, mapping of number onto a line becomes linear (the "logarithmic-to-linear shift," Siegler, Thompson, & Opfer, 2009). This logarithmic-to-linear shift has been observed across a wide variety of number-line tasks, including ones that vary the magnitude of numbers over several orders of magnitude (Dehaene et al., 2008; Opfer & Siegler, 2007; Siegler & Opfer, 2003; Thompson & Opfer, 2010), that present symbolic or non-symbolic numbers (Dehaene et al., 2008; Sasanguie, De Smedt, Defever, & Reynvoet, 2012; Sella, Berteletti, Lucangeli, & Zorzi, 2015), that present bounded or unbounded number lines (Kim & Opfer, 2017; Qin, Kim, & Opfer, 2017), and that provide instruction with or without anchors (Opfer, Thompson, & Kim, 2016).

Compressive mappings are also present in adults' estimates under certain circumstances—lack of formal education (Dehaene et al., 2008), attentional resources (Anobile et al., 2012; Cicchini et al., 2014; Dotan & Dehaene, 2016), or time (Dotan & Dehaene, 2013, 2016)—that arguably put adults on a more equal cognitive footing with children. Effects of formal education were found in a study that examined number-line estimation among the Amazonian Mundurucu without formal education and American adults (Dehaene et al., 2008). When logarithmic and linear regressors were tested, the logarithmic regressor better predicted Mundurucu adults' estimates than linear regressor, whereas the reverse was true for educated American adults—unless numbers were large and non-symbolic. Thus, it appears that logarithmic compression may be the default spatial-numeric mapping without schooling.

Logarithmically compressive mappings are also evident to some degree among educated adults if they are placed under attentional load. For example, when attentional resources were taxed in a dual task (number-line and another irrelevant task), adults' estimates had a strong logarithmic component (Anobile et al., 2012; Cicchini et al., 2014; Dotan & Dehaene, 2016). To quantify the degree of logarithmicity, Anobile et al. (2012) used a mixed log-linear model (MLLM; cf. Dehaene et al., 2008; Karolis, Iuculano, & Butterworth, 2011) that integrates logarithmic and linear functions:

$$y = a\left(\lambda \frac{U}{\ln(U)} \ln(x) + (1 - \lambda)x\right), \tag{1}$$

where y denotes the estimate to a given number x, a is a scaling factor that scales a represented magnitude of x, U the number at the right end of a number line, and λ an index of logarithmicity. If estimates are perfectly linear, λ equals 0, whereas λ is 1 for completely logarithmic estimates (Fig. 1B). Adults produce estimates with less compression ($\lambda = 0.11$ for adults in Cicchini et al., 2014) compared to young children ($\lambda = 0.77$ for children in Opfer et al., 2016). However, even educated adults showed greater logarithmicity ($\lambda = 0.38$) when attention was deprived by an additional task given with an estimation task (Cicchini et al., 2014).

Finally, logarithmic components are evident in the early processing stages of number-line estimation. Dotan and Dehaene (2013, 2016) tracked finger trajectories while adults were positioning a given number on a number line. In the study, fingers were observed to move first toward a position that was consistent with a log-scaled number position around 500–1050 ms and was later corrected to the direction to a more linear position of a number (Dotan & Dehaene, 2013). Taken together, findings from adult number-line estimation indicate that formal education provides the knowledge that allows top-down control of an inherent logarithmic bias, but this control that suppresses the logarithmic influence requires attention and time (but also see Dotan & Dehaene, 2016, for an alternative explanation for healthy adults).

Understanding the source of compressive mappings is both theoretically and practically important. Theoretically, it provides a fundamental understanding of an inherent mental number line that humans appear to share with nonhuman species (Nieder & Dehaene, 2009) and of basic quantitative skills, such as comparing, approximating, and projecting combinations of quantities (Opfer & Siegler, 2012). In addition, logarithmic compression in spatial-numeric mappings has been viewed as important evidence for numerical representation following the Weber-Fechner law, which describes the relationship between the physical magnitude and the subjective sensation (Siegler & Opfer, 2003). According to the psychophysical principle, the perceived magnitude increases as the logarithm of an actual magnitude. Practically, nonlinearity in number-to-space mapping also predicts several real world outcomes, including accuracy of memory for number (Thompson & Opfer, 2016; Thompson & Siegler, 2010), number categorization (Opfer & Thompson, 2008), dyscalculia (Geary et al., 2008), math scores (Ashcraft & Moore, 2012; Fazio, Bailey, Thompson, & Siegler, 2014; Gunderson et al., 2012; Kim & Opfer, 2017; Sasanguie, Göbel, Moll, Smets, & Reynvoet, 2013), and quality of economic utility judgments (Peters & Bjalkebring, 2015; Schley & Peters, 2014).

1.2. Sources of compression

1.2.1. Static logarithmic encoding

Why might early, untutored number-line estimates be logarithmically compressed? Two (non-mutually exclusive) explanations have been offered. The classical explanation is that early magnitude estimates are logarithmic due to an innate encoding system first proposed by Dehaene and Changeux (1993). In the neural network model, numerosity detectors that are inherently hardwired to respond selectively to a preferred numerosity, forming a Gaussian distribution of activity on a log scale. Because of the log-normal coding, overlaps among activations of numerosity detectors increase with number. For example, numerosity detectors that peak to 9 objects are more activated by 14 objects than by 4 objects despite the same distance between 9 vs. 14 and 9 vs. 4. Thus, this log-normal feature explains numerical performance obeying the Weber-Fechner law and underlying mechanisms of a logarithmic encoding of numbers (Dehaene & Changeux, 1993; Verguts & Fias, 2004; also see Gallistel & Gelman, 2000, for linear encoding with increasing variability).

Independent empirical support for a logarithmic encoding mechanism like that by Dehaene and Changeux (1993) came from a series of later findings by Nieder and colleagues (for a review, see Nieder, 2005). These investigators obtained single-cell recordings of neural activity as awake monkeys tracked the number of objects in a set during a delayed match-to-sample task. They found that neurons in the prefrontal cortex (PFC) and the primate analog of the human intraparietal sulcus (IPS) showed clear numerosity tuning, with peak firing for a preferred numerosity and decreasing activity as numerosity deviated from the preferred magnitude (Nieder & Merten, 2007). Critically, the tuning curves of these 'numerons' showed log-normal variability as modeled in Dehaene and Changeux (1993). Thus, collectively, the number-tuned neurons identified by Nieder and colleagues formed a physiological basis for the Weber-Fechner law.

1.2.2. Dynamic encoding

Against the logarithmic encoding account, an alternative view—the dynamic encoding hypothesis (Anobile et al., 2012; Cicchini et al., 2014)—is that compressive mapping can be attributed to a "central tendency of judgment," which is the tendency of estimates to be biased toward the mean of a stimulus distribution (Cross, 1973; Garner, 1953; Hollingworth, 1910; Jesteadt, Luce, & Green, 1977; see Gescheider, 1988, for review). For example, the magnitude of 15 is likely to be underestimated if preceded by 1 and overestimated if preceded by 29 as the central value of tested stimuli moves toward 1 or 29.

According to the dynamic encoding account, numbers are represented linearly with scalar variability (i.e., fuzzier and less confident representation for larger number). As a result, if estimating a larger number, responders rely more heavily on prior knowledge of previously encountered stimuli and integrate this information with the linear representation of the current magnitude. This dynamic process gives rise to the central tendency of judgment as the central values of the mental set (i.e., the prior) is updated online in the course of estimation (also see Anobile et al., 2012, for a fixed prior). From this perspective, logarithmicity in number-line estimation reflects the dynamic encoding mechanism, such that a current estimate is influenced by the number previously encountered, which tends to anchor the next estimate. Specifically, under a central tendency, numbers are underestimated if preceded by a smaller number trial, whereas numbers are overestimated if preceded by a larger number trial. Therefore, a logarithmic pattern of estimates on a number line does not come from a "static logarithmic transform," but emerges online from regression of estimates toward the mean of the tested distribution over trials.

Evidence in favor of the dynamic encoding hypothesis came from a study by Cicchini et al. (2014). They asked five adults to estimate the position of a non-symbolic number (a set of dots) on a 0–100 number line, with 9 unique numerosities tested on a total of 144 trials. The researchers used a dual-task paradigm to elicit more compression in mappings by boosting a central tendency. According to the dynamic encoding account, greater logarithmic compression was expected in a dual-task with additional attentional demands. This was because greater cognitive demands could increase response noise—i.e., Weber fraction in the likelihood—or reliance on the history of numbers previously experienced, which in turn leads to a greater central tendency (Anobile et al., 2012). Cicchini and colleagues examined the nonlinearity of estimates using the MLLM (Eq. (1)). Logarithmicity (λ) of estimates was greater in a dual task than that in a single number-line task, supporting the idea that the lack of attentional resources yields more logarithmic estimates. Furthermore, when regressed against numerosity from the previous trials, estimates on the current trial were under-estimated after smaller previous numerosity, whereas overestimated after larger previous numerosity in both single- and dual-task conditions. Compared to the single-task condition, under attentional overload, the inter-trial dependency was stronger along with greater logarithmicity.

Cicchini et al. (2014) further proposed a mathematical model to account for logarithmic patterns that might have resulted from the sequential dependencies. The model derived from a Kalman filter produces a current estimate based on the knowledge of previous trials while simultaneously updating the response distribution on the current trial. Cicchini et al. (2014) showed that the model reliably predicted compressive estimates as well as captured the inter-trial dependencies—i.e., underestimation after small trials and overestimation after larger trials—that cannot be simulated using the conventional MLLM. The integration model fitted estimates better than the MLLM in a dual-task condition, where sequential effects were stronger. On the other hand, in a single-task condition, in which estimates were more linear, the model did not predict estimates better than the MLLM.

Together, these results suggest that more attentional demand elicited greater logarithmicity *and* a stronger central tendency. From these findings, Cicchini et al. (2014) concluded that attentional load increased logarithmicity *because* it increased a central tendency. Based on the supposed cause-effect relation between dynamic encoding and logarithmicity, they claimed, "the strongest evidence for logarithmic coding [of number] was the logarithmic number line: Because that now has a more plausible explanation, there exists no evidence at all for logarithmic encoding of number in primate brains" (p. 7871).

1.3. How dynamic and logarithmic encodings might interact to produce a logarithmic-to-linear shift

Theoretically, logarithmic vs. dynamic encoding of numbers are not mutually exclusive—the nature of the default encoding is conceptually distinct from the nature of sequential effects. Ideally, each of these should be jointly modeled to check for potential interactions. In this section, we examine some potential interactions that might occur.

Our basic idea is that in a number-line task like the one in Fig. 1A, estimates can be viewed as ratings of the similarity of the target number (e.g., 15 dots) and endpoint (e.g., 30 dots). When the target and endpoint are highly similar (e.g., 29 and 30), the subject marks a position on the line that is close to the endpoint. When the target and endpoint are dissimilar (e.g., 1 and 30), the subject marks a position on the line that is far from the endpoint (for this interpretation of number-to-space mappings, see Cantlon, Cordes, Libertus, & Brannon, 2009). Because similarity ratings are highly context-dependent (Goldstone, Medin, & Halberstadt, 1997; Medin, Goldstone, & Gentner, 1993; Nosofsky, 1988; Parducci, 1965; Tversky, 1977), number-line estimates would be expected to be influenced by context as well (e.g., reference points, Holyoak, 1978).

In terms of sequential effects, the most important context effects concern the effect of similarity among pairs of compared items on later similarity judgments (Sjöberg & Thorslund, 1979; Tversky, 1977). A common finding in this research is that if participants had previously judged the similarity of low-similarity items (e.g., wasp-chicken), the similarity ratings of highly similar items (e.g., goose-chicken) increased; in contrast, after participants had rated the similarity of high-similarity items (e.g., sparrow-chicken), the similarity ratings of high-similarity items (e.g., goose-chicken) decreased (Sjöberg & Thorslund, 1979). In the psychophysical literature, these context effects are often considered to result from adjustment to an internal *response scale* over trials, which is known as the adaptation level (Helson, 1964; Parducci, 1965) or average response level (Petrov, 2008; Petrov & Anderson, 2005).

The same sequential effects on response scaling should be expected in the case of number-line estimation. If one had just judged the similarity of low-similarity items (e.g., 1 and 30), the similarity rating of 15 and 30 would be exaggerated (as on a compressive scale), but if one had just judged highly similar items (e.g., 29 and 30), the similarity rating of 15 and 30 would be minimized (as on an expansive scale). These context effects must also be weighted. If no memory trace is left by a previous trial (e.g., for trials too long before the current trial or in subjects with poor memory or attention), context cannot have an effect and the weight should be set to 0 when the weight is between 0 and 1 (Staddon, King, & Lockhead, 1980; Ward & Lockhead, 1970). In contrast, to the extent that the subject responds more to previous trials than the current trial, the weight would increase from 0.5 towards 1.

Thus, number-line estimates seem likely to be a product of three key components. The first component concerns the psychological scaling of numbers (as indexed by λ in Fig. 1B) before any other numbers have been presented. The next component concerns the differences between number pairs in successive trials (e.g., 15 and 30 vs. 29 and 30 when 15 is presented after 29). For number lines with a fixed upper bound, the differences between pairs reduce to the differences between numbers presented in previous and current trials (e.g., 29 − 15; $x_{i-1} - x_i$ in Eq. (4) below). In the case of presenting the same number repeatedly, $x_{i-1} - x_i$ is 0 (e.g., 29 − 29 = 0), but in the case of presenting different numbers, the component is used to readjust the response scale of numbers (λ in Eq. (4)). The third component concerns the weight of the sequential effect (*w* in Eqs. (2) and (4) below), which must be 0 on the first trial and is generally found to be equal to 0 on all trials other than the immediately preceding one (Gescheider, 1988).

To check the plausibility of this idea, we formulated a dynamic mixed log-linear model (D-MLLM) that incorporated the three components and simulated its performance under several conditions. The D-MLLM includes two major parts: a psychological scale of numbers that is either logarithmic or linear and a sequential dependency on the previous trial. It is formulated as follows:

$$y_i = (1 - w)p_i + w \cdot x_{i-1}, \quad i > 1, \tag{2}$$

where y_i indicates an estimate on *i* trial, p_i a magnitude encoding of an *i*-trial number, and *w* a weight of a number presented prior to the current trial (x_{i-1}) on the estimate of the next number tested.[1] The weight parameter is between 0 and 1, such that it can reflect the relative influence of sequential effects and representation of current trial. In our simulations below, a conservative estimate of the weight parameter (*w*) comes from the strength of the central tendency of judgment (which integrates sequential effects over all trials), but the weight component more broadly reflects the relative influence of sequential effects and representation. This weight parameter produces dynamic changes in the response scaling that would mimic changes in the psychological scaling, given by Eq. (3):

$$p_i = a\left(\lambda_i \frac{U}{\ln(U)} \ln(x_i) + (1 - \lambda_i)x_i\right). \tag{3}$$

p_i is similar to the MLLM (Eq. (1)) except that the representation of numerical magnitude is updated from trial to trial based on how similar the current number is to the upper bound in a context of the previous trial:

$$\lambda_i = \lambda_{i-1}(1 - w((U - x_i) - (U - x_{i-1}))) \tag{4}$$

$$= \lambda_{i-1}(1 - w(x_{i-1} - x_i)).$$

[1] Note that Cicchini et al. (2014) modeled sequential effects using previous estimates (y_{i-1}) instead of previous stimuli (x_{i-1}), assuming that previous estimates would be the best replacement for what responders remember for previous stimuli. However, their empirical test used previous stimuli (not responses) to test for sequential effects, arguing that the central tendency of judgment can be directly probed with preceding stimuli. For simplicity, our current model also uses a stimulus that is presented before the current trial to model and test sequential dependencies, but note that it is also possible for the current trial to rely on previous estimates or both of previous estimates and stimuli (Mori & Ward, 1995; Petrov & Anderson, 2005).

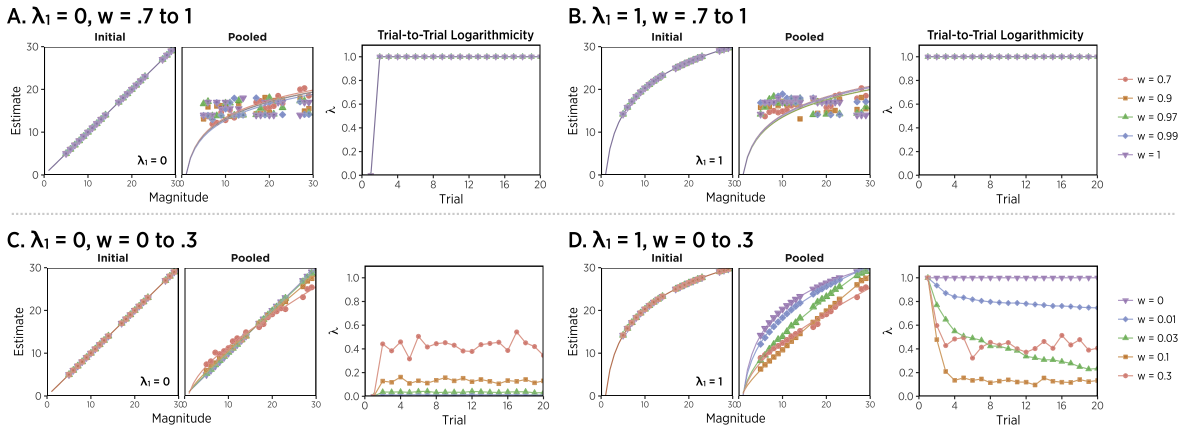

Fig. 2. Simulations of 300 virtual participants using the D-MLLM. First-trial estimates, median estimates collapsed over trials, and trial-to trial changes in the compression index are presented across different initial logarithmicity (λ_1) and weight parameters (w). For the simulations, scaling factor a equaled 1, and 20 numbers used in the present paper were used in random order.

In Eq. (4), the logarithmicity on i trial (λ_i) is between 0 and 1 and adjusted from the previous logarithmicity (λ_{i-1}) based on the difference between previous (x_{i-1}) and current (x_i) stimuli, weighted by the parameter w. Therefore, in a case, where 15 is estimated on a 0–30 number line, the logarithmicity *decreases* if the current trial follows a larger trial (e.g., 29) that exaggerates the dissimilarity of 15 to 30; the logarithmicity of the current trial *increases* after a trial of a smaller number (e.g., 1) that makes the similarity between 15 and 30 more salient. This postulate provides interesting predictions for order effects in estimates. If participants are given stimuli in an ascending order, estimates will be more logarithmic, whereas they will be more linear if stimuli are given in an descending order (Fig. A1). Critically, there is no weight on the first trial—i.e., $w = 0$—because there have been no previous trials before the first trial. Thus, on the first trial, the D-MLLM is exactly the same as the MLLM.

A noise-free simulation of D-MLLM is presented in Fig. 2. As in Cicchini et al. (2014)'s argument, when initial (first-trial) estimates were simulated as completely linear ($\lambda_1 = 0$) and the weight of the previous stimulus had been set high ($w = 0.7$ to 1), median estimates pooled over all trials were perfectly compressive ($\lambda = 1$), and the logarithmicity index increased after just a few trials (Fig. 2A). Similarly, when initial estimates were set as completely logarithmic ($\lambda_1 = 1$) and the weight of the previous stimulus was high ($w = 0.7$ to 1), median estimates remained highly compressive ($\lambda = 1$), though without trial-to-trial changes in logarithmicity (Fig. 2B). In this sense, a strong central tendency of judgment would have a compressive effect on median estimates (i.e., pooled over all trials).

Simulations from the same model also depict what happens when the effect of the previous trial is smaller ($w = 0$ to 0.3). When initial estimates were set to be linear (Fig. 2C, $\lambda_1 = 0$), median estimates barely increased in compression for smaller w's ($\lambda = 0.01$ to 0.11 when $w = 0$ to 0.1) and more noticeably for a larger w ($\lambda = 0.38$ if $w = 0.3$). But if initial estimates were completely logarithmic ($\lambda_1 = 1$), the logarithmic component of median estimates fell fast with w's that were between 0.03 and 0.3 ($\lambda = 0.14$ to 0.46), with most of the change occurring over the first 10 trials (Fig. 2D), whereas median estimates stayed logarithmic over time for zero or almost zero w's ($\lambda = 0.78$ to 1 if $w = 0$ to 0.01). Thus, initially logarithmic estimates could either remain so when the weight of previous trials is high or give rise to a logarithmic-to-linear shift when the weight is low (e.g., 0.03 to 0.3).

Theoretically, this interaction between a high logarithmicity component and small sequential effect is very interesting because it can address some surprising developmental differences, such as age- (and educational-) differences in estimating non-symbolic number (e.g., Dehaene et al., 2008; Qin et al., 2017; Sella et al., 2015; this manuscript, Study 4). In the D-MLLM model, those who appear to represent number linearly on the number-line task (like educated older children and adults) may nevertheless encode number logarithmically, but remember and take advantage of previous stimuli to linearize their responses. On the other hand, children with limited memory resources would not be expected to exploit metric information about sequential numbers and thus to rely on logarithmic encoding alone. In this way, the D-MLLM allows us to distinguish between age-related changes in dynamic response scaling from age-related changes in numerical representations.

2. Overview of studies 1–4

The current studies revisit whether any causal relation exists between sequential effects and compressive estimates of numbers, as well as to test predictions from D-MLLM. To test whether dynamic encoding is the sole source of logarithmic estimates, Studies 1–4 examined participants' initial estimates in a number-line task. On the first trial, no sequential effect is possible because no previous number has been encountered. Study 2 addressed the same issue using a same-numerosity design, where participants were repeatedly tested on the same numerosity. Thus, like first-trial information in Studies 1–4, inter-trial differences were non-existent in Study 2, so estimates would be expected to remain either logarithmic or linear.

To examine how dynamic processes might interact with logarithmic encoding over the course of estimation, Studies 1–4 also examined logarithmicity of estimates on a trial-by-trial basis, such that the predicted dynamic effects of D-MLLM could be tested. If the initial encoding mode was logarithmic and interacted with a dynamic process (as in our D-MLLM simulation), a logarithmic-to-linear shift would be expected in Studies 1, 3, and 4, but not Study 2. These same issues were explored in adults' estimates of non-symbolic numbers alone (Studies 1, 2, and 3), as well as in children's and adults' estimates of non-symbolic and symbolic numbers (Study 4).

3. Study 1

In Study 1, we sought to replicate sequential effects with a larger number of adult participants (n = 40) and more numerosities (20 numerosities between 0 and 30) than previously tested (Cicchini et al., 2014; n = 5 and 9 numerosities between 0 and 100). The increase in sample size was necessary to obtain the statistical power necessary to make meaningful inferences about single-trial estimates. In Cicchini et al. (2014), numerosities as small as 2 and 3 were used with large numerosities beyond the subitizable range. Considering such a short presentation of an array of dots (240 ms), it is highly likely that small numerosities were easily tracked, whereas magnitude of large numerosities, like 67, were estimated with difficulty. Moreover, it is possible that distinct numerical systems were applied in approximating magnitude, depending on the subitizability of numerosity (Feigenson, Dehaene, & Spelke, 2004). The authors also reported that there were no effects of previous trials for small, subitizable numerosities (Cicchini et al., 2014). In the current study, only numerosities larger than 4 were included, such that the same numerical system would be in effect for all stimuli. We also expected that estimation of the larger number of numerosities on a smaller scale would allow us to examine dynamic effects in number-line estimation more thoroughly.

3.1. Method

3.1.1. Participants

Forty undergraduate students at the Ohio State University participated in the study (M = 20.55 years, SD = 3.08 years, 19 male students). All the participants had normal or corrected-to-normal vision. They received credit in return for their participation.

3.1.2. Procedure and design

Participants were given a non-symbolic number-line task, in which they were shown a set of dots (5–29) on a computer screen and asked to estimate the number of dots by mouse-clicking a position on a line flanked by 0 dots on the left and 30 dots on the right (Fig. 1A). On each trial, a set of to-be-estimated dots were shown briefly (240 ms) and immediately followed by a random-noise mask, employed to prevent counting. This procedure continued for 3 blocks of 20 trials. The to-be-estimated numerosities were chosen to sample non-subitizable numbers evenly: 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 17, 18, 19, 20, 21, 22, 23, 27, 28, 29. Compared to the previous study, in which only 9 numerosities on a 0–100 scale were used (Cicchini et al., 2014), the current study included most of the numerosities between 0 and 30 except for those that were subitizable or might serve as landmarks. By doing so, mappings of all possible numerical values onto a number line and their trial-by-trial changes were examined more exhaustively. The order of numerosities presented was determined by a balanced Latin square, such that each numerosity was presented to two participants on each of the 20 trials in a block. The size of dots was controlled on half the trials: the dot size was equal to the size of the 30 dots at the right end for half the trials, while the total surface area covered by dots was the same as the aggregate of the surface of the right-end dots on the other half trials. After instructions, participants started the task with no practice trials or feedback of any kind.

3.2. Results

If dynamic encoding were the only cause of logarithmic numerosity estimates, the central tendency of judgment should accompany logarithmic compression in estimates. Thus, if estimates are not logarithmic, but completely linear, the central tendency of judgment would not be expected. To examine this, we first tested if median estimates were logarithmically compressive. The median positions of participants' estimates were regressed against the number of dots presented. For this purpose, we used the MLLM (Eq. (1)). As shown in Fig. 3A.1, collapsing over all the 60 trials, median estimates were completely linear ($\lambda = 0$), replicating findings from schooled participants in Dehaene et al. (2008).

Weights of the logarithmicity (λ) parameter were next tracked on a trial-by-trial basis using median estimates of each numerosity collapsing over participants on each trial (a scatter plot for each trial is presented in Fig. A2). Against the idea that dynamic encoding is solely responsible for logarithmic estimates, a large logarithmic component was evident on the first trial ($\lambda = 0.85$), but decreased drastically to the last trial ($\lambda = 0.09$) (Table 1 and Fig. 3A.2). To test this observation statistically, we regressed the logarithmicity values on trial number. Consistent with nominal values, logarithmic index values decreased with trial number somewhat exponentially, $\tau = 3.14$ in $y = e^{-x/\tau}$ and $b = -1.08$, $p < .01$ in a regression with log-scaled trial number and log-scaled logarithmicity.

To check whether logarithmic-to-linear changes came merely from practice reducing noise, averaged standard deviations of estimates were computed on every trial and regressed against trial number. The amount of noise did not decrease over time ($b = 0.003$, $p > .05$). These results indicate that logarithmic number-line estimates do not require any dynamic response to previous stimulus distributions; rather, they are strongly present on the first trial. Additionally, if participants' estimates were subject to any inter-trial dependencies, the dynamic processes appeared to have linearization effects that contrast with the effects of central tendency that bring about compression.

We next examined whether any central tendency was present. To test this, we first followed Cicchini et al. (2014)'s analytic

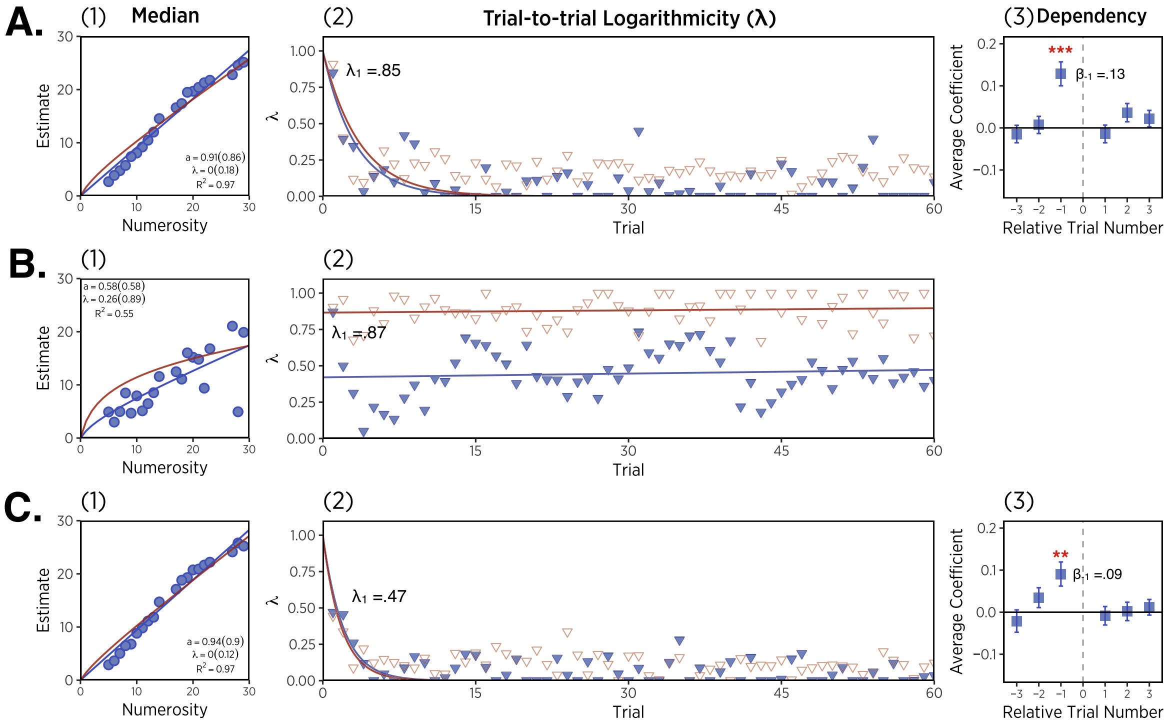

Fig. 3. Actual data (blue) and D-MLLM predictions (red) in Study 1–3 (A-C respectively): (1) median estimates of actual and simulated data. Parameter values of the best-fitting models are shown for real data along with those for simulated data in parentheses. (2) Trial-to-trial logarithmicity of real and simulated estimates. (3) Mean coefficients as a function of relative trial number in actual estimates. Error bars indicate standard errors. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Logarithmic component (λ) in Study 1–4.

	Group	First trial	Last trial	All trials
Study 1				
	Adult	0.85	0.09	0
Study 2				
	Adult	0.87	0.40	0.26
Study 3				
	Adult	0.47	0.04	0
Study 4				
Non-symbolic	Child	1	1	1
	Adult	0.70	0.18	0.11
Symbolic	Child	0.36	0.32	0.15
	Adult	0	0	0

procedure, in which errors for each trial were categorized into 3 groups based on magnitude of previous stimuli (larger by 6, similar, or smaller by 6). We found that, consistent with a central tendency, when a set of dots was presented after a larger set, participants tended to overestimate the number of dots. Also, after a smaller set, participants tended to underestimate the number of dots. Thus, as reported in Cicchini et al. (2014), the average error of estimates differed reliably as a function of magnitude of numerosities prior to current ones ($F(2, 20) = 23.32$, $p < .001$).

To test for dynamic effects more systematically, we again regressed current responses on numerosities on the past and the future trials. Fig. 4 shows the dependency of current estimates of two sample numerosities on the immediately prior trials (-1 trials). If current responses were more dependent on previous numbers, estimates of current numbers would be more underestimated after small previous trials and more overestimated after large previous trials, resulting in a greater regression coefficient. Consistent with Cicchini et al. (2014), there was a tendency for estimates of large numbers to be more dependent on previous trials (e.g., the greater regression coefficient for 22 than 5 in Fig. 4). The same regression analysis was applied to current estimates of all 20 stimuli on all the possible past and future trials (i.e., from -59 to $+59$ trials) to ensure that there was no artifactual dependency on future trials (Cicchini et al., 2014). We next conducted bootstrap significance tests on the regression coefficients of each trial. As found in Cicchini et al. (2014), regression weights were the greatest and significantly different from zero in the immediately previous trial (average

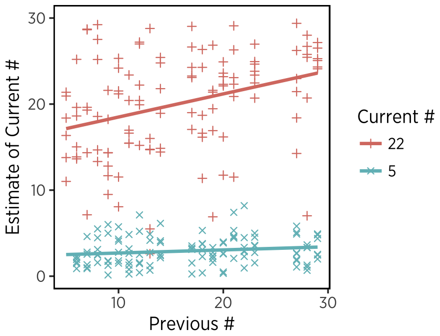

Fig. 4. Estimates of two sample numerosities (22 and 5) as a function of numerosity on immediately previous trials (−1 trials).

$\beta = 0.13$, $p < .001$) (Fig. 3A.3). Thus, a central tendency of judgment was evident overall, but (collapsing over these trials) very little compression was observed.

To test our alternative model of dynamic processing, 300-subject responses to the same stimuli were simulated using the D-MLLM. The scaling factor from median estimates pooled over all trials was adopted for the scaling factor in simulation ($a = 0.91$ in Fig. 3A.1). The logarithmicity component for the first trial in adults' real estimates was used for the initial logarithmicity ($\lambda_1 = 0.85$). For the dependency parameter, the mean coefficient (average β) of the −1 trial obtained by regressing current estimates on immediately prior stimuli was employed ($w = 0.13$). Simulated estimates were corrupted with errors randomly generated from a Gaussian distribution. The variability of the error distribution was determined using the averaged standard deviation of actual estimates in current Studies (e.g., $SD = 4.87$ in Study 1). Simulated estimates that were smaller than the lower bound (0) or larger than the upper bound (30) were adjusted to be at the closest bound.

Identical analyses conducted on adults' estimates were done on the simulated data. The red line in Fig. 3A.1 shows the best-fitting MLLM to median estimates from simulations. The medians of simulated responses collapsing across all trials were somewhat linear, exhibiting relatively small logarithmicity ($\lambda = 0.18$) as in actual estimates and as in Cicchini et al. (2014) 's single-task condition. The medians of simulated estimates fit the medians of actual estimates well ($R^2 = 0.93$). When logarithmic compression was analyzed trial by trial, simulations also presented a decreasing trend in logarithmicity ($\tau = 3.61$, and $b = -0.18$, $p < .05$ in a log-log regression) (Fig. 3A.2). The simulation results are consistent with findings of a logarithmic-to-linear shift in actual estimates. In simulation, the decrease in logarithmicity was mainly driven by the interaction between a weak dynamic encoding effect and an initially strong logarithmic component (as in Fig. 2D).

Lastly, sequential dependencies on previous trials were examined. Regressed on the magnitudes of numerosities immediately prior to the current trials, more underestimation of current trials was evident after small trials, $F(2, 20) = 107.9$, $p < .001$. When regressed on stimuli on relative trials (−59 to +59 trials), current estimates were considerably dependent on −1 trials although the dependency was greater than actual data (average $\beta = 0.18$, $p < .001$). Taken together, the simulation demonstrated that even if a central tendency is evident as in estimates, dynamic encoding might have given rise to a linearizing process that is distinct from and can overweigh a central tendency over the course of a task.

4. Study 2

In Study 1, a strong logarithmic component was observed on the first trial, before any dynamic process could give rise to results. In Study 2, we tested for a logarithmic component over trials, but without varying inter-trial numeric differences, by presenting the same numerosity repeatedly to a given participant. If the effect of dynamic updates from previous trials increased compression, the amount of compression would be expected to remain low from the first to last trial in Study 2, and the amount of compression observed in Study 1 would be greater than Study 2. In contrast, if dynamic processes lead to linearization as in the D-MLLM, the amount of compression in estimates would be expected to remain high from the first to the last trial, and the amount of compression in Study 1 would be less than Study 2.

4.1. Methods

4.1.1. Participants

In this Study, 40 undergraduate students from the Ohio State University were recruited ($M = 18.84$ years, $SD = 3.15$ years, 14 male students). They were awarded credit toward a psychology class in return for their participation.

4.1.2. Procedure and design

Study 2 was identical to Study 1. It consisted of 3 blocks of 20 trials in which numerosity was displayed for only 240 ms. The only difference between the two Studies was that in this Study participants estimated the same numerosity repeatedly for all trials. Each numerosity was estimated sixty times by two participants who were randomly assigned to the number. The location of each dot in a set was determined randomly on each trial so that a display of an array of dots differed trial by trial. On the other hand, the size of dots was not controlled and always the same as the 30-dot size at the right end. By presenting the same sized and the same number of dots on every trial, we attempted to keep subjective experience of a numerosity identical to avoid any changes in perceived magnitude of the same numerosity that can result from irrelevant stimulus dimensions (e.g., the size of dots; Gebuis & Reynvoet, 2012).

4.2. Results

The MLLM was first fit to median estimates that collapsed over all 60 trials. The median estimates showed slightly greater compression ($\lambda = 0.26$) than those in Study 1 ($\lambda = 0$), where central tendency effects were present (Fig. 3B.1). Additionally, the average logarithmicity for a given trial in Study 2 ($M = 0.45$, $SD = 0.17$) was considerably greater than that in Study 1 ($M = 0.10$, $SD = 0.15$), $p < .0001$ in a randomization test. These empirical results are not at all consistent with the model depicting that a central tendency of judgment is required for compression.

Trial-by-trial analyses on the logarithmic components also yielded evidence against the dynamic encoding explanation for compression, revealing that the first-trial estimates ($\lambda = 0.87$) were more logarithmic than those for later trials and as logarithmic as the estimates on the first trial in Study 1, where first trial compression was the same as in this Study, $p = .88$ in a randomization test. Moreover, unlike Study 1, no reduction in compression was observed from trial to trial, $b = 0.13$, $p > .05$ in a log-log transformed regression and $b = 0.001$, $p > .05$ in a linear regression (Table 1, Fig. 3B.2, and Fig. A3). Thus, at least qualitatively, both the differences in overall compression and trial-to-trial changes are more consistent with the predictions of the D-MLLM than the dynamic encoding account.

To test predictions of the D-MLLM, responses of 300 virtual participants were simulated in the same manner as Study 1. For the two parameter values from actual data (i.e., first-trial λ and a scaling factor of median estimates) were used for λ_1 and a in simulation ($\lambda_1 = 0.87$, $a = 0.58$). Because there was no possible dependency on the previous trial in the same-numerosity task, the weight parameter was set to be 0 ($w = 0$). When collapsed over trials, the medians of simulated estimates were more logarithmic than actual responses ($\lambda = 0.89$, Fig. 3B.1), and the over-compression tendency resulted in a misfitting of simulated medians to actual medians, $R^2 = 0.26$. Consistent with the findings in medians, trial-by-trial logarithmicity was greater than that of actual estimates (Simulated: $M = 0.88$, $SD = 0.10$), $p < .001$ in a randomization test. As shown in real data, however, a trial-by-trial analysis revealed that the first trials were logarithmic ($\lambda_1 = 0.90$), and the logarithmicity stayed constantly high over time, $b = 0.001$, $p > .05$ in a linear regression (Fig. 3B.2). Although the simulation produced more compressive responses than those observed in human participants, the simulation accurately captured the fact that estimates could stay logarithmic from the beginning to the end of a number-line task.

5. Study 3

In Study 3, we examined whether estimates would be less compressive when subjects had more time to encode numerosity. Study 3 was identical to Study 1 except for a longer display of a stimulus (2000 ms). Previous research had shown that longer exposure to stimuli improves numerosity discriminability (Kaswan & Young, 1963) and line-length estimation (Crawford, Huttenlocher, & Engebretson, 2000). For example, Crawford et al. (2000) showed that estimates of line length were more biased toward the center of a stimulus distribution in the short display condition. According to a Bayesian approach by Huttenlocher and colleagues, if participants have an inexact memory for a stimulus due to constraints on encoding time or cognitive resources, they adjust estimates more to the central value of the category, leading to a stronger central tendency of judgment (Huttenlocher, Hedges, & Duncan, 1991). However, it still remains unknown whether logarithmicity in number-to-space mappings is dependent on display time. Also, if linear mapping is not automatic, but requires time, attentional resources, and better memories for stimuli, a longer display may yield more linear responses as a result (Anobile et al., 2012; Dotan & Dehaene, 2013, 2016; Huttenlocher et al., 1991). This issue is potentially important in interpreting previous results. In Cicchini et al. (2014)'s study, a stimulus was on the screen for only 240 ms, which is extremely short compared to typical studies with number-line tasks (e.g., Anobile, Stievano, & Burr, 2013; Fazio et al., 2014).

5.1. Methods

5.1.1. Participants

In Study 3, 41 undergraduate students at the Ohio State University participated ($M = 20.05$ years, $SD = 1.69$ years; 16 male students). One participant who misunderstood and positioned the same location for all numerosities was excluded from analyses.

5.1.2. Procedure and design

Study 3 followed the same procedure as Study 1 except that participants were exposed to a stimulus for 2000 ms.

5.2. Results

The logarithmic component (λ) for overall estimates was computed as in Studies 1 and 2. Consistent with Study 1, median estimates exhibited no compression ($\lambda = 0$) (Fig. 3C.1). Estimates from the first to last trials were tracked, and values of the logarithmic index were highest on the first trial ($\lambda = 0.47$) and exponentially decreased to 0.04 with trial number, $\tau = 1.99$ and $b = -0.80$, $p < .05$ in a log-log regression (Table 1, Fig. 3C.2, and Fig. A4). Again, the logarithmic-to-linear transition did not result from response noise ($b = -0.006$, $p > .05$). Thus, as in Study 1, results indicate that a central tendency was not required for logarithmically compressed mapping.

To test whether the logarithmicity was affected by longer stimulus exposures, a randomization test was performed on first-trial estimates of Study 1 and 3. Even though the logarithmicity on the first trial for Study 3 was lower in value, it did not significantly differ from that of Study 1 ($p = .39$). In the same manner, the mean difference of 60-trial logarithmicity values from the two Studies was analyzed. There was no statistical difference in the average logarithmicity of 60 trials in Study 1 ($M = 0.10$, $SD = 0.15$) and Study 3 ($M = 0.08$, $SD = 0.10$) either, $p = .31$. Therefore, the effect of a longer stimulus exposure on compression in mappings was not evident (though might be if display time were long enough for counting).

Sequential effects were again evident in Study 3. The mean errors of current responses significantly changed depending on immediately previous trials, $F(2, 20) = 26.49$, $p < .001$. The result implies that there were similar systematic biases in responses: Overestimation after a large-numerosity trial and underestimation after a small-numerosity trial. This trend was again observed when estimates were regressed against magnitude of the past and the future trials. First, regressed on magnitude of the -1 trial, the current estimates of 5 numerosities were significantly predicted by previous stimuli ($p < .05$ for numerosity 17, 19, 20, 22, and 29). Second, when estimates on current trials were regressed as a function of numerosity of relative trials, magnitudes of stimuli immediately prior to current trials affected current responses significantly (average $\beta = 0.09$, $p < .01$) (Fig. 3C.3).

Simulation was next conducted in the identical way in Study 1 and 2, using actual initial responses for the logarithmicity, median estimates for a scaling factor, and the average β of -1 trial for the weight parameter ($\lambda_1 = 0.47$, $a = 0.94$, $w = 0.09$). First, even though the median responses in simulation were slightly more logarithmic ($\lambda = 0.12$ for medians, Fig. 3C.1), the simulation showed good model-fitting to adults' median estimates ($R^2 = 0.95$). The trial-by-trial logarithmicity components presented a decrease over time, $\tau = 1.82$ and $b = -0.43$, $p < .05$ in a log-log regression (Fig. 3C.2). Moreover, as found in actual estimates, current trials were over- or under-estimated depending on the magnitude categories of immediately previous trials (i.e., smaller, similar, or large), $F(2, 20) = 51.06$, $p < .001$. When current estimates were regressed against trials presented in the past and the future, the weight of trials immediately prior to the current ones was significantly different from zero (average $\beta = 0.15$, $p < .001$).

In summary, the effects of display duration on compression in number-line estimates were examined using a longer display time than in Study 1. Consistent with findings in Study 1, the logarithmic pattern was evident on the first trial, and the compressed estimates became quickly linearized after the first trials. Even though a set of dots stayed on the screen for 2000 ms, the longer display of numerosity affected logarithmicity neither on the first trial nor over all the 60 trials in mappings. Again, the experimental and simulation results point to the logarithmic-to-linear shifts during number-to-space mappings that may result from incorporating metric information about differences between successively tested numbers.

6. Study 4

In Study 4, we examined potential interactions of sequential effects with numerical formats and age. Study 4 followed a more conventional design in number-line studies, which typically uses symbolic numbers and presents each stimulus only once (Siegler & Booth, 2004; Siegler & Opfer, 2003). Although Cicchini and colleagues claimed that logarithmic compression was the result of dynamic encoding, their arguments were based only on adults' estimates of non-symbolic numbers (Anobile et al., 2012; but see Cicchini et al., 2014 for a brief summary on children's inter-trial dependency). However, considering a large body of studies showing developmental changes in estimates of symbolic numbers, it is important to explore relations between logarithmicity and sequential effects in estimates of symbolic and non-symbolic numbers by both children and adults. In this study, we examined kindergartners between the ages of 5 and 6 years. Given that children at those ages would be at chance levels when required to hold memories for more than 3 items serially presented (i.e., 2-back tasks) (Ciesielski, Lesnik, Savoy, Grant, & Ahlfors, 2006; Kwon, Reiss, & Menon, 2002), we expected children to show little to no between-trial dependencies in number-line tasks, where 3 numbers (a given number and two ends) are presented on each trial.

6.1. Methods

6.1.1. Participants

Seventy 5- to 6-year-old children were recruited from upper-middle- or middle-class schools in Columbus ($M = 6.02$ years, $SD = 0.45$ years, 37 male students) and 80 undergraduate students from the Ohio State University ($M = 19.75$ years, $SD = 2.37$ years, 45 male students). We chose the age ranges for child participants because children between 5 and 6 have much less experience with linear number-lines than older children, thereby helping to isolate the influence of schooling (if any) on numerosity estimation. Children were given a sticker in return for their participation, whereas adults received credit toward a psychology course.

6.1.2. Procedure and design

Participants completed both symbolic and non-symbolic number-line tasks that were given in a random order. Study 4 followed the similar procedure to Study 3 except for the smaller number of trials, with an additional number-line task in symbolic number. Each task included only one block of 20 trials, which randomly presented 20 stimuli only once. The same 20 numbers used in Study 1 and 3 were used and presented for 2000 ms for both numerosity and Arabic number conditions in the Study. In the symbolic condition, an Arabic numeral was estimated on a number line with numeral 0 and 30 at the ends. There was no practice or feedback provided for both age groups as in a typical setting (Laski & Siegler, 2007; Opfer & Siegler, 2007; Opfer & Thompson, 2008).

6.2. Results

6.2.1. Estimation of non-symbolic numbers

To test dynamic effects on logarithmic mappings, logarithmicity of median estimates and trial-by-trial responses were analyzed with the MLLM. First, when collapsing over all trials, children's median responses were completely logarithmic (λ = 1), whereas adults' median estimates were more linear (λ = 0.11), a λ value that was identical to that observed by Cicchini et al. (2014) in the single task condition. (Table 1 and Fig. 5A.1 and B.1). A randomization test revealed that the logarithmic component of children's median estimates was significantly greater than that of adults' ($p < .001$).

Next, logarithmicity values were computed on a trial-by-trial basis to examine how compression changed over time. Children indeed showed constant logarithmic patterns in numerosity estimation on the first, the last trial, and nearly every trial in between (λ = 1 for the first trial, λ = 1 for the last trial) (Table 1, Fig. 5A.2, and Fig. A5). In contrast, there was a trend for the logarithmic components of adults' estimates to decrease over trials (λ = 0.70 for the first trial, λ = 0.18 for the last trial; Fig. 5B.2), and the 20 logarithmic components of adults' responses were significantly lower than those of children's estimates (children: $M = 0.97$, $SD = 0.09$; adults: $M = 0.21$, $SD = 0.19$), $p < .0001$, randomization test. However, unlike Studies 1 and 3, the decrease in adults' logarithmicity scores was not significant, $\tau = 3.09$ and $b = -0.39$, $p = .47$ in a log-log transformed regression and $b = -0.012$, $p = .11$ in a linear regression, presumably due to the smaller number of trials in this study. Consistent with this conjecture, the logarithmicity scores of adults in Study 4 did not differ from those of the first 20 trials of Study 3 ($M = 0.12$, $SD = 0.14$), $p = .11$.

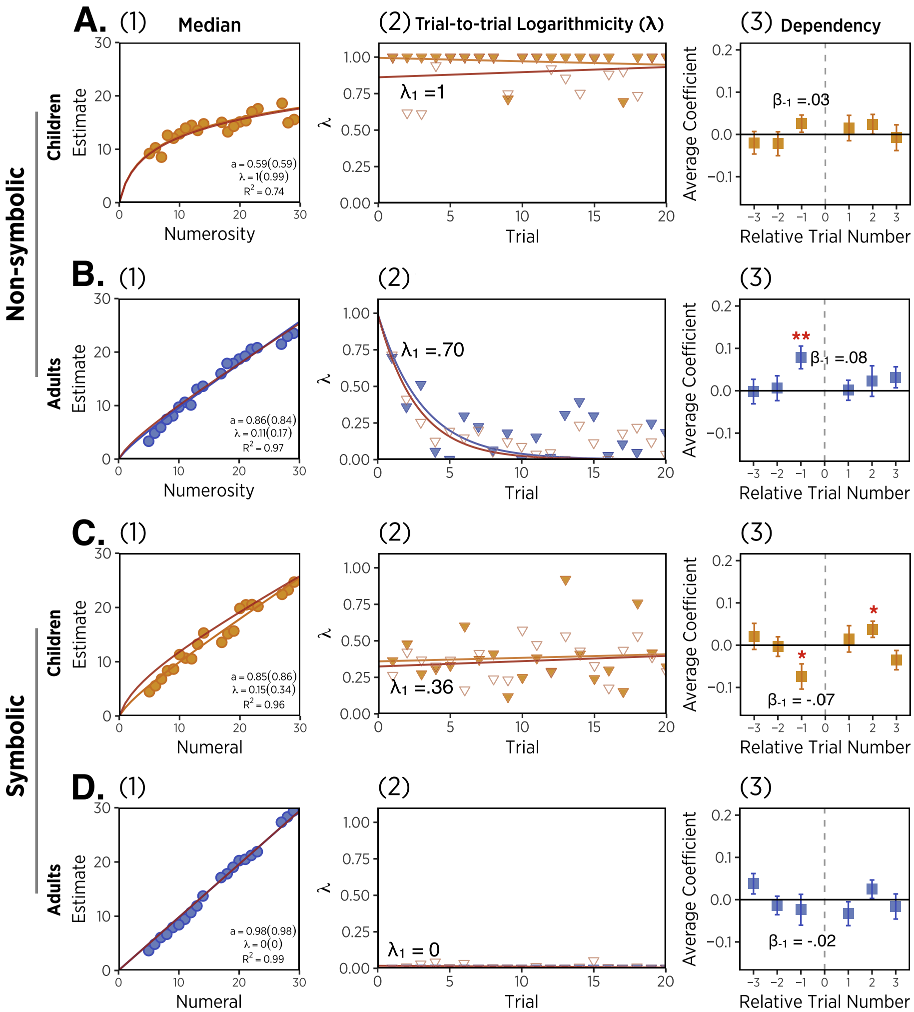

Fig. 5. Children's (yellow) and adults' (blue) actual data and D-MLLM predictions (red) in Study 4 (A and B for non-symbolic, C and D for symbolic condition): (1) median estimates of actual and simulated data. Parameter values of the best-fitting models are shown for real data along with those for simulated data in parentheses. (2) Trial-to-trial logarithmicity of real and simulated estimates. (3) Mean coefficients as a function of relative trial number in actual estimates. Error bars indicate standard errors. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

These results narrowly replicate previous findings demonstrating age differences in number-line estimation (Siegler & Opfer, 2003). However, on the first trial, the logarithmicity values of adults' and children's estimates (0.70 for adults vs. 1 for children) did not differ from each other, $p = .30$ in a randomization test, and were much closer in value than those on the last trial (0.18 for adults vs. 1 for children). The difference on the first and last trials reflects the fact that there was a trend for the logarithmicity of adults' estimates to decrease from trial to trial. Overall, these results were not consistent with the dynamic encoding explanation for compressed mapping: Compression was again evident on trial 1, before any dynamic encoding mechanism could have an influence. Nor were these results consistent with the idea that the representation of numerical magnitudes changes much from childhood to adulthood: On the first trial of numerosity estimation, adults' estimates were as logarithmic as children's. Rather, trial-to-trial analyses suggest that the native representation of numerosity is logarithmic, but this impression can be supervened (at least in adults) by encountering numerosities that differ from trial to trial.

To test whether this effect of trial-to-trial changes might have occurred via a central tendency, response errors were again grouped by the magnitude of the previous trial. The sequential effects were not found in children's responses, but adult mappings changed significantly as a function of the previous trial ($F(2, 30) = 6.60$, $p < .01$). Again, a numerosity was more underestimated when the previous number was smaller than when it was similar or larger in adults. Unlike Cicchini et al. (2014) or Study 1 and 3, however, overestimation after the large-numerosity trials was not observed. To examine this more closely, estimates were regressed on magnitude of trials given in the past and the future. For children, there was no special relationship between the current response and trials given before or after the current trial, average $\beta = 0.03$, $p = .20$ for -1 trials (Fig. 5A.3). For adults, the dependencies of the current mapping were found on the immediately previous trial, average $\beta = 0.08$, $p < .01$ for -1 trials (Fig. 5B.3). Thus, as in Studies 1 and 3, we observed that the central tendency of judgment co-occurred with a decrease in estimation compression.

To check whether the co-occurrence of a central tendency of judgment and a decline in estimation compression was consistent with our D-MLLM model, we generated estimates of 300 virtual children and adults. Again, the logarithmic index from the first trial and a scaling factor from median estimates were adopted in simulation ($\lambda_1 = 1$, $a = 0.59$ for children, $\lambda_1 = 0.70$, $a = 0.86$ for adults). Given no sequential dependencies in children, we posited the dependency parameter as zero for children, whereas the average β from -1 trials was used for the weight in adults ($w = 0$ for children and 0.08 for adults). When collapsed over all 20 trials, simulated medians for children were logarithmic as in real responses ($\lambda = 0.99$, Fig. 5A.1) with a rather poor fitting to actual median data ($R^2 = 0.70$). Medians from virtual adults were actually quite close to real responses ($\lambda = 0.17$, Fig. 5B.1), $R^2 = 0.96$. Trial-to-trial changes yielded similar results. The D-MLLM correctly predicted no changes in trial-to-trial logarithmicity, $b = 0.004$, $p > .05$ in a linear regression in children (whose actual estimates were logarithmic on every trial) (Fig. 5A.2). The D-MLLM simulated a decreasing pattern in logarithmicity for adults, $\tau = 2.70$, $b = -1.01$, $p < .05$ in a log-log regression (Fig. 5B.2).

When sequential effects in simulated data were tested, the mean errors of current responses did not change as a function of previous numerosity in children's simulated estimates, $F(2, 30) = 0.228$, $p > .05$. However, unlike children's real estimates, their simulated responses exhibited a significant reliance on the -1 trials when regressed as a function of relative trials, average $\beta = 0.03$, $p = .04$. For adults, previous numerosities had significant effects on the mean errors, $F(2, 30) = 3.591$, $p < .05$. Regressed against numerosities on relatively past and future trials, simulated estimates were also dependent on the -1 trials, showing the mean regression coefficient that was significantly greater than zero, average $\beta = 0.09$, $p < .001$. Again, the findings in simulations suggest dynamic encoding as a source of linearity in number-line estimates by shifting the internal scaling over time.

6.2.2. Estimation of symbolic numbers

It is important to examine logarithmic compression and inter-trial dependencies in mappings of symbolic numbers since they do not rely on perceptual processing as heavily as mappings of non-symbolic numbers would. Recent work has shown that symbolic numbers are represented more accurately and linearly than numerosities in later development once associations between numerical symbols and their magnitudes are acquired. Among Mundurucu participants, for example, educated ones mapped numbers on a number line linearly when stimuli were given in Portuguese, while estimating non-symbolic numbers on a log scale (Dehaene et al., 2008). The same patterns were reported in estimates by children: Greater linearity in the symbolic conditions than in the non-symbolic conditions. (Sasanguie et al., 2012; Sella et al., 2015).

The similar tendency was also found in median and trial-by-trial estimates in the current study. Compared to the median responses for non-symbolic numbers, the median estimates of symbolic numbers were more linear for both groups ($\lambda = 0.15$ for children, $\lambda = 0$ for adults). Adults' median estimates were perfectly linear, and the difference between the two groups in logarithmicity was significant, $p < .01$, randomization test (Table 1, Fig. 5C.1 and D.1, and Fig. A6).

Children's and adults' estimates on each trial were fit by the MLLM to compute trial-by-trial logarithmicity. As expected, the logarithmic component value of children's symbolic estimates on the first trial ($\lambda = 0.36$) was lower than that of children's non-symbolic estimates in value. The linear patterns in children's responses stayed somewhat constant across trials, $b = 0.003$, $p = .75$ in a linear regression (Fig. 5C.2). For adults, estimates of symbolic numbers were so accurate in the first place that there was no compression detected on any trials ($\lambda = 0$ for all trials) (Fig. 5D.2). When the difference between two groups in the logarithmic components tracked trial by trial was evaluated, adults' performance in symbolic numbers ($M = 0$, $SD = 0$) was more linear than children's ($M = 0.38$, $SD = 0.19$), $p < .0001$. These findings again replicate previous studies that have shown the developmental improvement in number-line estimation (Siegler & Opfer, 2003; Siegler et al., 2009). The logarithmic-to-linear changes over trials observed in numerosity mappings by adults, however, were not found in estimation of symbolic numbers.

The dynamic encoding effects on symbolic number estimation were examined with averaged response errors by magnitude of the previous trial. As opposed to prediction by the dynamic encoding hypothesis, children have a tendency to underestimate the current number more when it followed a larger number trial, but magnitude of the previous trial did not show significant effects on the current responses, $p > .05$. For adults, although the current stimulus tended to be more underestimated when given after a smaller number, the magnitude of −1 trial was not a significant main effect, $p > .05$. Next, the current estimates were regressed against stimuli from −19 to +19 trial. Inconsistent with the central tendency, children's estimates were negatively related to previous numbers, average $\beta = -0.07$, $p = .02$ (Fig. 5C.3). An unexpected dependency was also evident between the current and a future trial—i.e., +2 trials, average $\beta = 0.04$, $p = .04$. The retroactive predictability of future trials on the current estimates does not make logical sense and might be a false positive. Interestingly, the dependencies on the immediately previous trial found in Study 1 and 3 could not be found in adults, $p > .05$ (Fig. 5D.3).

Estimates of two groups were regenerated with the D-MLLM. Simulations were conducted in the identical way in Study 1–3. For children, a weight parameter was set to 0, the lowest value of w in the model ($\lambda_1 = 0.36$, $a = 0.85$, $w = 0$). For adults, a zero weight parameter was used as well ($\lambda_1 = 0$, $a = 0.98$, $w = 0$). Simulated medians exhibited greater compression ($\lambda = 0.34$, Fig. 5C.1), but fit the medians of children's real estimates well, $R^2 = 0.90$. Despite the absence of dynamic effects, simulations reproduced perfectly linear medians in adults ($\lambda = 0$, Fig. 5D.1), $R^2 = 0.99$. Analyses for trial-to-trial logarithmicity on simulated data exhibited no decrease in logarithmic components over time in children, $b = 0.01$, $p > .05$ in a linear regression (Fig. 5C.2). In adults' virtual responses, zero logarithmicity appeared to remain constant over trials, $b = -0.001$, $p > .05$ in a linear regression (Fig. 5D.2).

To test dynamic encoding effects in simulations, simulated responses were regressed against previous trials. For children, simulations accurately modeled no main effect of the −1 trials on the mean errors as in actual estimates, and no inter-trial dependency on the −1 trials, $ps > .05$. For adults whose simulations entailed no compression to linearize over trials, the simulated estimates showed sequential effects in neither averaged errors nor regression coefficients as a function of past and future trials, $ps > .05$.

In summary, the effects of dynamic processes were tested with estimates of non-symbolic and symbolic numbers by young children and adults. It replicated the previous studies showing the developmental improvement in number-line estimation: Logarithmic compression was greater in children's overall estimates than in adults'. Consistent with literature on numerical cognition, estimates of symbolic numbers were more accurate and linear than those of non-symbolic numbers. However, trial-by-trial analyses revealed that adults' mapping of numerosity was as logarithmic as children's on the first trial, where there was no previous trial that could possibly affect the current response. The compression diminished over the course of estimates in adults, whereas the compression was constantly high in children. Moreover, the simulation could successfully mimic a non-decreasing and a decreasing trend in logarithmic compression in children and adults. The results suggest that representation of non-symbolic numbers for educated adults may be logarithmic as for children and other species, but may be corrected to be linear with consideration of magnitude of the previous stimulus over trials.

7. Discussion

In this paper, we sought to identify the role of dynamic encoding in compressive number-to-space mapping. According to the dynamic encoding account, the sequential dependency between previous and current trials is the sole cause of compression in number-line estimation (Anobile et al., 2012; Cicchini et al., 2014). Here, we sought to investigate the alternative possibility that dynamic response scaling (i.e., changes in the function relating the internal representation to the response) has an interactive effect with the psychological scaling of number (i.e., the function relating the external stimulus to the internal representation), much as is observed in similarity ratings and category judgments (Helson, 1964; Parducci, 1965; Petrov & Anderson, 2005). To test this, we revised our static model of psychological scaling (MLLM) to make use of sequential dependencies in a dynamic model (D-MLLM). Predictions from the D-MLLM were then tested in various circumstances and across ages.

7.1. Compressive mapping is the default, no dynamic encoding required

Against the dynamic encoding account, our results indicate that a central tendency of judgment is not required for logarithmic compression in numerical estimates. If dynamic updates were necessary for logarithmic estimates, the logarithmic component would be expected to be near zero prior to any previous numbers encountered (i.e., the first trial). Across Studies 1–4, however, we found that the logarithmic component of adults' estimates on the first trial was much more compressive, with the observed logarithmicity ranging from 0.47 to 0.87. These results are not consistent with the idea that the psychological scaling of numerosity is linear.

To test this idea still further, we also simulated data from a dynamic encoding model (Cicchini et al., 2014), in which numerosities are scaled linearly with scalar variability (Fig. A7). For the first trial without any dynamic effects, the model assumes linear representation of numerosity and entails Gaussian errors with the zero mean and the scalar *SD* (i.e., $k \cdot x^{\alpha}$, increasing variability unless α equals 0). As in Cicchini et al. (2014), the power-law parameter α and the constant k were obtained directly from our data using a simple regression and model fitting to the data respectively (see Cicchini et al., 2014, for details). Simulations of the dynamic encoding model also supported the idea that the very first estimates of non-symbolic numbers would be quite linear in adults ($\lambda = 0$ to 0.10 in Study 1–4) (see first trials in Fig. A7). Thus, even a model assuming linear scaling with scalar variability would predict much lower levels of compression (0–0.10) than were actually observed (0.47–0.87).

Additionally, sequential effects did not even coincide with compression. In Study 1 and 3, for example, sequential effects were evident after trial 1, yet the central tendency did not appear to elicit logarithmic responding. Rather, as predicted by our D-MLLM, median estimates pooled over all trials were perfectly linear, and the logarithmicity of estimates *decreased* rather than increased from trial to trial. In contrast, the dynamic encoding model failed to predict the log-to-linear shifts found in adults' estimates of numerosity (Fig. A7). For example, in Study 1, estimates simulated from the dynamic encoding account were linear on the first trial ($\lambda = 0.10$) and on following trials ($M = 0.12$, $SD = 0.04$ in λ's) without any significant shifts in logarithmicity ($b = 0.001$, $p = .19$ when logarithmicity was regressed against trial number). This result is consistent with the idea that the default perception of numerosity obeys the Weber-Fechner law, but is not at all consistent with the idea that compression requires dynamic encoding.

Single-number estimation in Study 2, which made it possible to separate effects of repeatedly estimating over trials from effects of viewing different numbers on different trials, was also inconsistent with the dynamic encoding hypothesis. According to the dynamic encoding account, single-number estimation lacks any informative prior trials; as a result estimates would be expected to stay linear from the first to last trials. Simulation of Cicchini et al. (2014)'s model also predicted constantly linear estimates over trials ($\lambda = 0.01$ and 0.03 for the initial and last trial respectively) (Fig. A7.B). Against this prediction, however, the logarithmicity of adults' actual estimates was greater than that of Study 1 and 3, where sequential effects were significant. Further, estimating a single numerosity repeatedly neither evoked the logarithmic-to-linear shifts that were found in Studies 1, 3, and 4, nor did the procedure reduce estimation variability from trial to trial. In fact, the amount of noise in estimates tended to increase from trial to trial, $b = 0.02$, $p < .001$. Again, these results would not be expected if the logarithmicity of numerosity estimates required a dynamic encoding mechanism, but this pattern was better predicted by our D-MLLM.

This compressive mapping in adults is surprisingly consistent with an earlier finding by Krueger (1982). Krueger (1982) was the first to attempt to separate dynamic processes from magnitude representation. To control for effects of previous trials, he adopted a single-trial design, in which participants estimated the numerosity of an X array only once, and which is thus comparable to the first trials in our studies. He found that single-trial estimates were still compressive ($\beta = 0.83$ of a power function). Our results extend this finding to a single-number design (Study 2), where estimates were fairly compressive on the first trial and every trial thereafter. Together, these results indicate that default encoding of number without any dynamic components is compressive.

7.2. Age-related changes in numerosity estimation: what develops?

Studies of number-line estimation typically find age- or education-related decreases in the logarithmicity (or compressiveness) of numerosity estimates (Dehaene et al., 2008; Qin et al., 2017; Sasanguie et al., 2012; Sella et al., 2015). This finding immediately raises the question of whether the perception of numerosity changes or something else. The present findings suggest that the psychological scaling of numerosity probably does *not* change much with age. Instead, age-differences in the linearity of numerosity estimates may reflect age-differences in encoding time and responsiveness to inter-trial differences in stimuli.

Several observations from our studies suggest that this view of development is correct. First, in Studies 1–4, the logarithmicity of adults' estimates on the first trial (i.e., before any sequential effects were possible) were 0.85 (Study 1), 0.87 (Study 2), 0.47 (Study 3), and 0.70 (Study 4), with compression associated with a 240-ms display time (Studies 1 and 2) being somewhat higher than with a 2000-ms display time (Studies 3 and 4). Logarithmicity of children's estimates on the first trial was 1 in Study 4, suggesting that age-related increases in the speed of numerosity encoding (Ratcliff, Love, Thompson, & Opfer, 2012; Thompson, Ratcliff, & McKoon, 2016) may play some role in increasing the linearity of numerosity estimates.

Second, age-differences in the logarithmicity of numerosity estimates were much larger in later trials than in earlier trials. In Study 4 (Fig. 5A-B.2), the logarithmicity parameters for adults and children were fairly close (0.70 vs. 1) on the first trial, but it dropped exponentially in adults but not in children. Thus, as effects of inter-trial differences could accumulate, developmental differences emerged, such that in the last ten trials adults' estimates were almost perfectly linear and children's estimates were almost perfectly logarithmic. Thus, whatever is causing differences between children's and adults' pattern of estimates must be emerging on a trial-to-trial basis, not as a function of aging. More exactly, these trial-wise decreases in logarithmicity occurred in adults only when they were asked to estimate different numbers on different trials (as in Study 1, 3, and 4). Asked to estimate the same number repeatedly (in Study 2), adults' estimates were stable. Thus, differences between children and adults were not present at first, emerged quickly over trials, but only when inter-trial differences in numeric stimuli were present.

All three of these findings were predicted by our D-MLLM model, which helps us to interpret what might be changing over development. In the model, an estimate depends on the relative weight of (1) the number that a subject sees in a given trial and (2) its difference from previous trials, much like the sequential effects observed in similarity ratings. Weighting previous trials *more* than current ones is obviously non-adaptive, which should lead us to expect the weight of sequential effects to be less than 0.5 (as in Fig. 2D instead of Fig. 2B). Empirically, the weight of sequential effects is also quite low, about 0.07 to 0.13 in adults (Cicchini et al., 2014, Fig. 3C; this paper, Fig. 3A.3, Fig. 3C.3, and Fig. 5B.3). In simulations of D-MLLM (Fig. 2D), even these small differences would be expected to generate differences in trial-to-trial changes in logarithmicity, such that even a weight that is half that of adults' would generate much less change.

Among young children, these inter-trial effects would almost certainly have to be attenuated or eliminated. Our expectation that children would not rely much (possibly at all) on previous trials comes from developmental differences on n-back tasks, which present subjects with a continuous series of a simple stimulus (e.g., letter O), and in which subjects are asked on each trial i whether the same stimulus was presented on trial i-n. In our task, each trial presented 3 numbers, with a various time interval between trials but never less than 2 s. For a 2-back task of even shorter duration than 2 s, growth models suggest that kindergartners would perform at chance levels (Ciesielski et al., 2006; Kwon et al., 2002), suggesting that they have no memory of $i-3$ number, whereas adults perform at more than 95% accuracy. At the very least, then, we should expect children to be much less likely than adults to be subject to inter-trial dependencies, and our observed weights were indeed not different from zero.

Interestingly, there was great similarity of trial-to-trial logarithmicity between children (Study 4) and adults who estimated the same numerosity (Study 2). The results are in line with a developmental study by Sullivan and Barner (2014), in which both linearity in estimates and the percent of responses relying on previous trials increased with age and memory supports (e.g., allowing children to see where they had marked all previous locations on the number line). Moreover, they found that individuals' degrees of reliance on previous trials predicted fits of a linear function. Taken together, there might be no developmental difference in representing the number of a set: Even educated adults might encode numbers logarithmically in the beginning, just like young children. However, by encoding more quickly and by remembering and using previous trials to adjust estimates on new trials, adults greatly improve the accuracy of their estimates.

7.3. Estimation of symbolic and non-symbolic value

Estimates of the value of symbolic numbers have also been reported to show age-related decreases in logarithmicity or compression (Barth & Paladino, 2011; Berteletti, Lucangeli, Piazza, Dehaene, & Zorzi, 2010; Heine et al., 2010; Laski & Siegler, 2007; Laski & Yu, 2014; Opfer & Martens, 2012; Opfer & Siegler, 2007; Opfer, Siegler, & Young, 2011; Opfer et al., 2016; Sasanguie et al., 2012; Siegler & Booth, 2004; Siegler & Opfer, 2003; Slusser, Santiago, & Barth, 2013; Thompson & Opfer, 2008; White & Szűcs, 2012). Could these changes, too, simply reflect developmental differences in responsiveness to previous stimuli?

Theoretically, we might expect the development of numerosity and symbolic number to be quite similar. Representation of numerosity seems fundamental for understanding symbolic numbers because the meaning of symbolic numbers is learned in part by associating numeric symbols with the quantities for which they stand (Dehaene, 2007; Opfer & Siegler, 2012). Consistent with this idea, logarithmic compression (λ) in estimates of symbolic and non-symbolic numbers in Study 4 was positively correlated in children ($r(68) = 0.32, p < .01$). That is, children who generated the most logarithmic estimates in numerosity estimation also generated the most logarithmic estimates in symbolic numeric estimation. Thus, we might expect a similar developmental trajectory.

Unlike numerosity estimation, however, estimates of symbolic numbers revealed developmental differences in numeric mapping in the absence of any sequential effects from prior trials. In Study 4, adults produced perfectly linear estimates from the first to the last trials, with zero inter-trial dependency (i.e., $w = 0$ and $\lambda = 0$). Also, children estimated Arabic numerals more logarithmically than adults from the first trial though the last, and also did not show valid inter-trial dependencies either (i.e., $w = 0$ but with high λ). Given no sequential dependencies in either age group and large age-differences in λ from the first through the final trial, the differences in logarithmic compression between children and adults seem to reflect actual developmental improvement in representing the value of numeric symbols.

Although developmental differences in estimates of symbolic number may not require sequential dependencies (as developmental differences in numerosity estimation did), one could imagine creating a situation where children would expect a certain sequence of numbers, be able to remember the previous number given, and use the difference in numbers to adjust their response scaling. One such case would be when children are given numbers in a counting sequence, either ascending or descending. In this case, D-MLLM could prove useful in generating predictions about the resulting pattern of symbolic number estimates (see, for example, Fig. A1).

7.4. Other characterizations of compressive mappings

In this paper, we have largely used logarithmicity (λ) from the mixed log-linear model to describe the degree of compression in children's estimates. Another interpretation of compressive mappings in number-line estimation is that estimates do not exactly follow the Weber-Fechner law, but Stevens' power law (Stevens, 1975). Based on Stevens' power law, some researchers have recently argued that compressive estimation stems from the lack of cognitive strategies, such as proportional reasoning (Barth & Paladino, 2011; Hollands & Dyre, 2000; Rouder & Geary, 2014; Slusser et al., 2013; Spence, 1990). According to the account, number-line estimates become compressive if responders rely only on one reference point (the lower end). Otherwise, estimation patterns may entail one or two cycles if estimates are made using two (the lower and upper ends) or three reference points (two ends and the midpoint) respectively.

Elsewhere, we have argued against the cyclic-power function approach due to the model complexity (Kim & Opfer, 2017; Opfer et al., 2011, 2016; Qin et al., 2017), but the simple power model seemed to be a straightforward equivalent to our MLLM. To examine this, we applied Stevens' power model ($y = a \cdot x^{\beta}$) to adults' estimates of numerosity in current studies. In Stevens' power model, the model parameter β is smaller than 1 as estimates are more compressive, whereas the parameter β approaches toward 1 as estimates become more linear. When fitting the model to estimates trial by trial, we found first estimates were compressive ($\beta = 0.45$ in Study 1; $\beta = 0.45$ in Study 2; $\beta = 0.64$ in Study 3; $\beta = 0.48$ in Study 4). We next regressed the trial-to-trial exponent β*'s* against trial number (e.g., 1–60 in Study 1–3 and 1–20 in Study 4). If there were shifts from compressive to linear during estimation, the parameter β would increase and get closer to 1 over trials. There were indeed considerable increases in the exponent β in all Studies except for Study 2, where the same numerosity was repeatedly estimated ($b = 0.003$, $p < .01$ in Study 1; $b = -0.001$, $p = .19$ in Study 2; $b = 0.002$, $p < .05$ in Study 3; $b = 0.009$, $p = .06$ in Study 4). Similar results were obtained using the absolute values between the trial-to-trial β*'s* and 1 (i.e., $|1 - \beta|$) as a non-linearity measure ($b = -0.001$, $p = .06$ in Study 1; $b = 0.001$, $p = .50$ in Study 2; $b = -0.002$, $p < .01$ in Study 3; $b = -0.01$, $p = .06$ in Study 4). The results are consistent with our finding that compression decreased in the course of a number-line task, especially in which different numerosities were estimated. Although Stevens' power law is very similar to the MLLM, however, it isn't immediately clear how to extend the model to incorporate a dynamic component (but see Cross, 1973, for a power function that takes biases of prior trials into account).

Still another characterization of compressive mappings employs a rational analysis approach. A recent study, for example, has shown that logarithmic patterns in number-line estimates may be the result of an internal scaling for the frequency distribution of number use—i.e., the need probability (Piantadosi, 2016). According to the author, in numerosity estimation, numerosity is mapped onto internal representation that is biased by noise. Estimators are thought to efficiently minimize the noise between numbers and the representation of numbers by adjusting their responses according to how often each number must be represented. It has been found that logarithmic compression can arise if the probability of representing number n is proportional to something similar to $1/n^2$. Although this account may not commit to the psychophysical assumptions of the logarithmic encoding account, it is not clear how it accounts for the present findings that estimates are initially logarithmic, but linearized over trials ***within a task***. Given that need probabilities are computed using a vast amount of data (e.g., frequencies of numbers that baboons encounter in the wild or of number words appearing on a corpus of books) (Piantadosi, 2016; Piantadosi & Cantlon, 2017), only a few encounters of different numbers would not seem likely to change the internal scaling of number. Possibly future work on the dynamics and development of numerosity estimation can elucidate the role of frequency information, which the present approach has not attempted to address.

8. Conclusions

Mapping number to space is not a static process, but changes with various dynamic components within a task and even within a single trial (Cicchini et al., 2014; Dotan & Dehaene, 2016). The present article sought to examine the effects of between-trial dynamics on number-to-space mapping in order to separate developmental differences in dynamic processes from developmental differences in the psychological representation of number. Our findings suggest that age differences in numerosity estimation do not reflect developmental differences in the psychological representation of number, but differences in dynamic processes. In contrast, age differences in symbolic number estimation reflect developmental differences in the psychological representation, not differences in dynamic processes.

Acknowledgments

The research is supported in part by IES grant R305A160295. We would like to thank the staff, parents, and children of Columbus Preparatory Academy, Our Lady of Bethlehem School and Childcare, and Wickliffe Progressive Elementary School for their generosity and support. We also thank Sungkyun Cho, Ariel Lindner, Marina Peeva, and Yiwan Wang for their assistance as well as Nora Newcombe, David Landy, and two anonymous reviewers for their valuable comments.

Appendix A

See Figs. A1–A7.

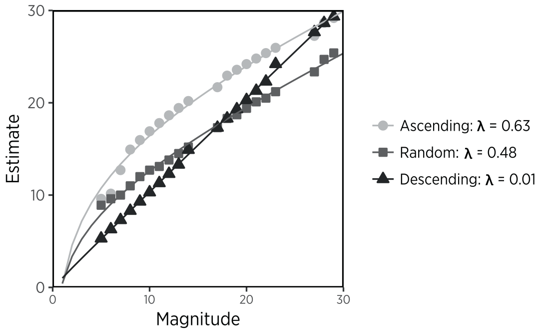

Fig. A1. Simulations of the D-MLLM (a = 1, λ_1 = 0.5, and w = 0.3). Median estimates pooled over all trials are presented as a function of stimulus order (i.e., ascending, random, and descending orders). Numbers used in the present studies were used.

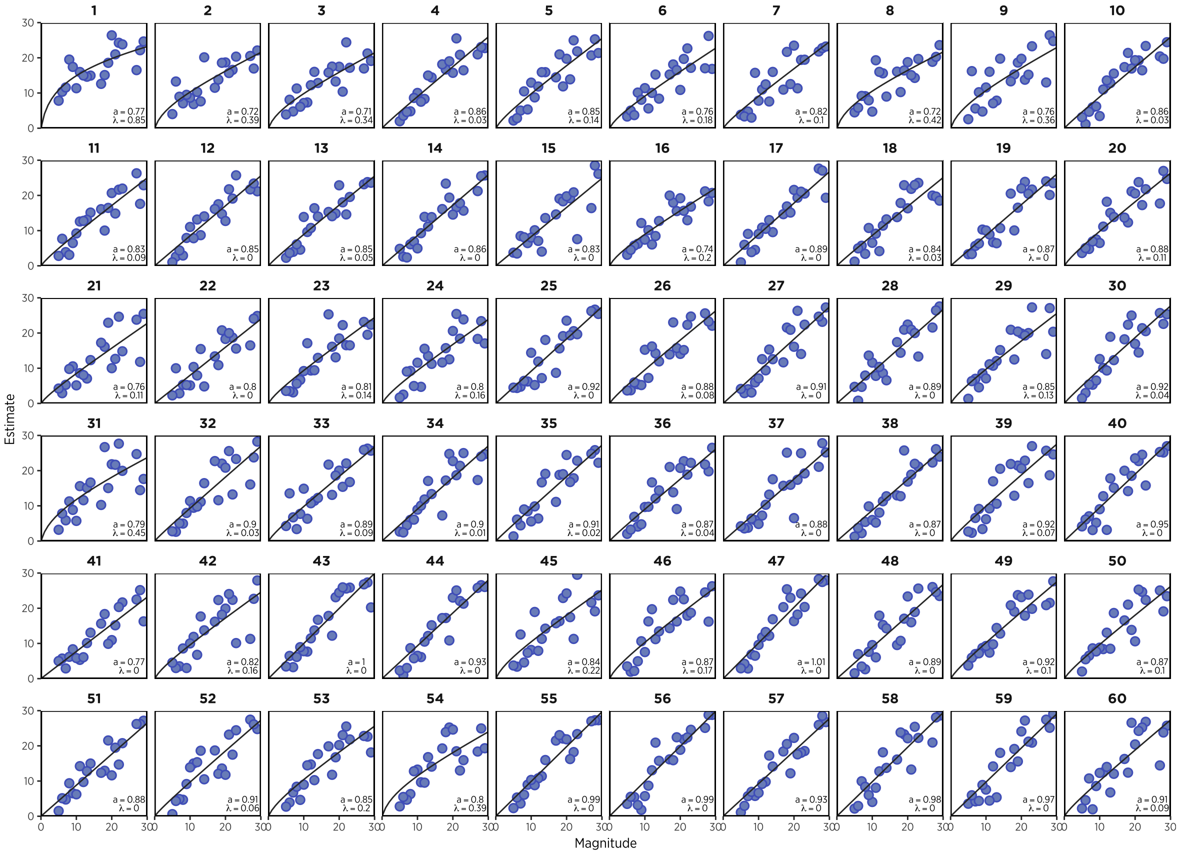

Fig. A2. Median estimates from the first to the last trials in Study 1.

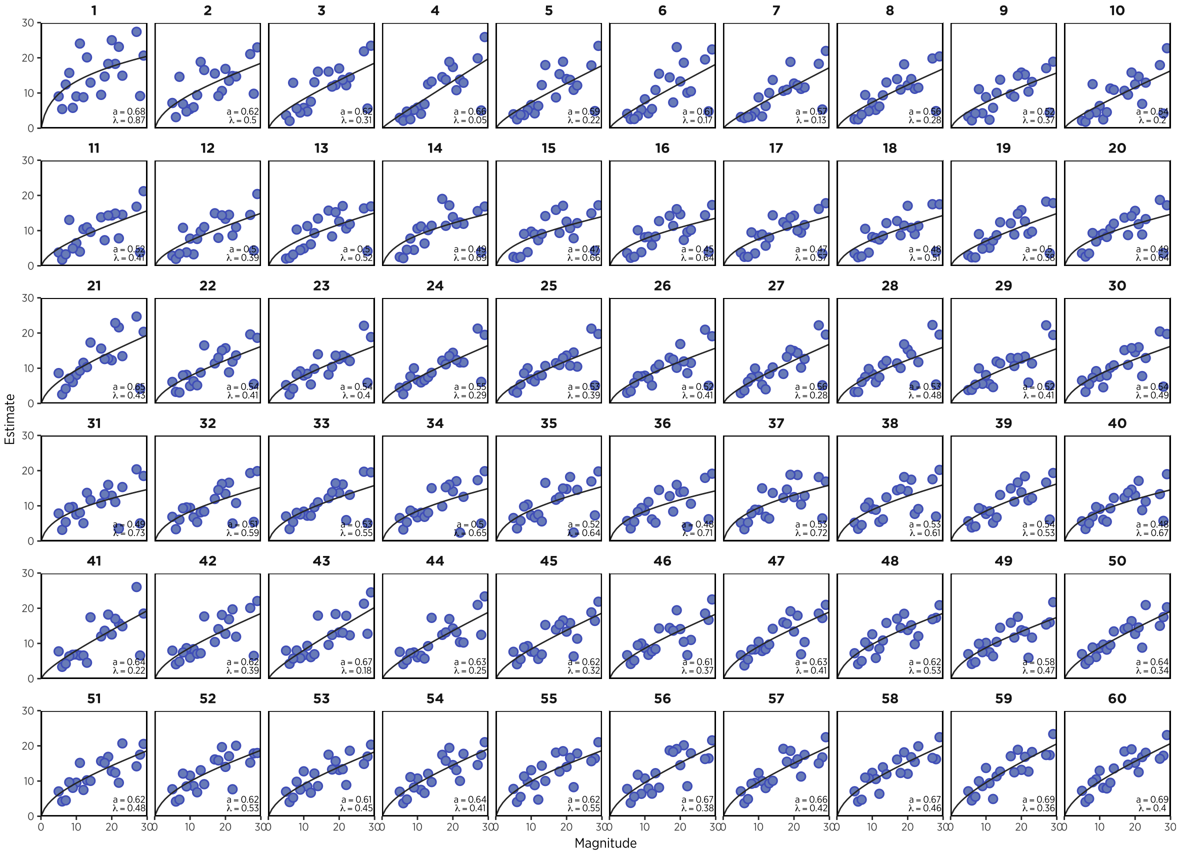

Fig. A3. Median estimates from the first to the last trials in Study 2.

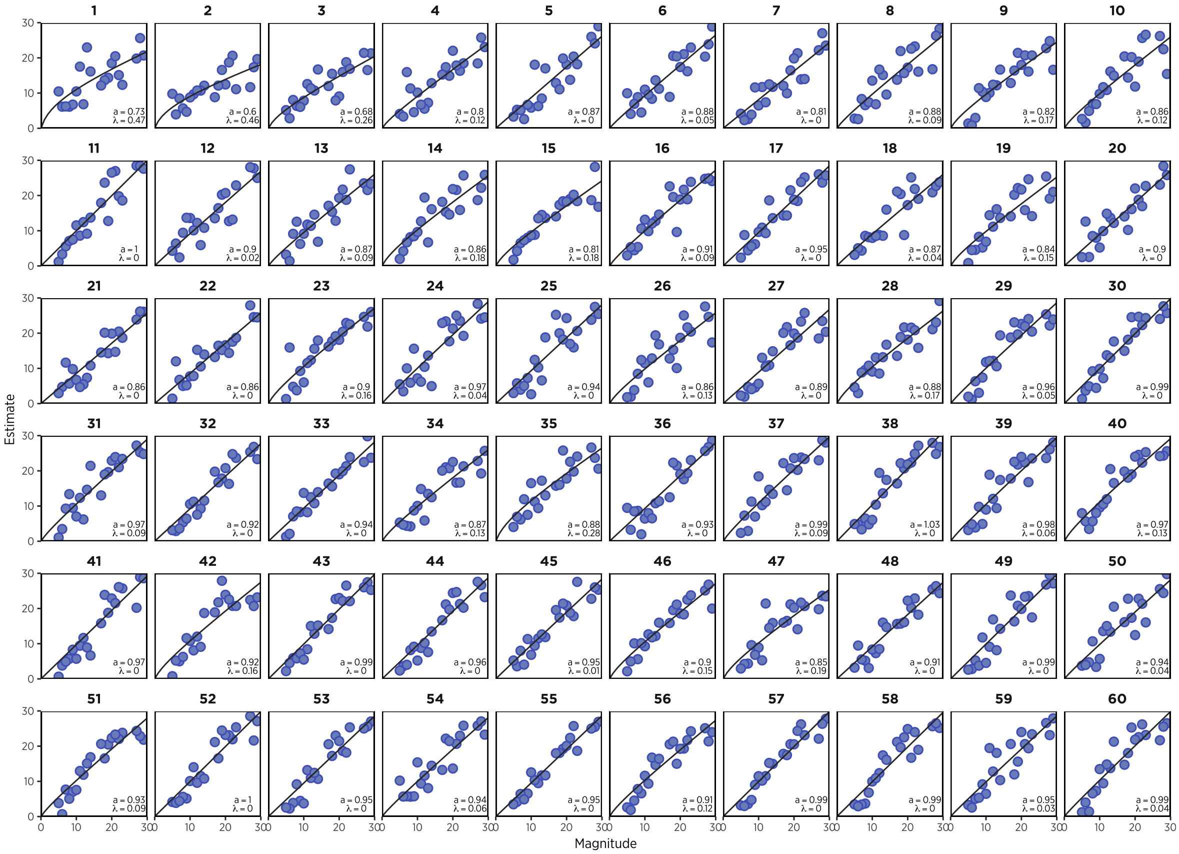

Fig. A4. Median estimates from the first to the last trials in Study 3.

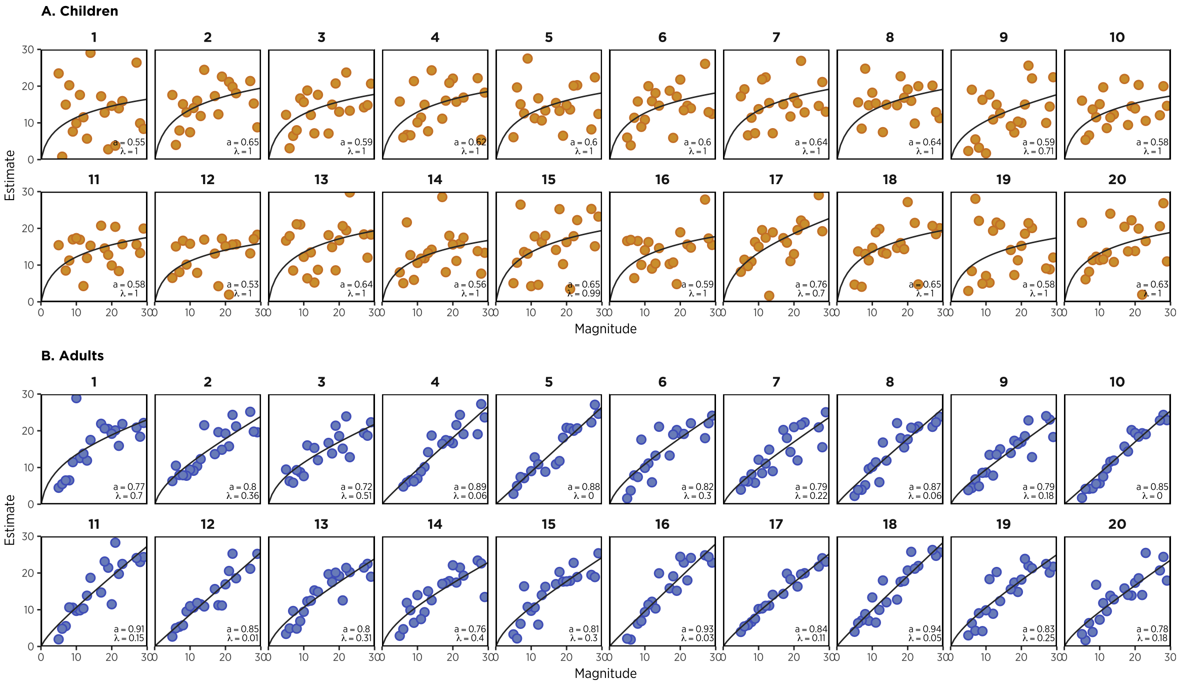

Fig. A5. Children's (A) and adults' (B) median estimates of non-symbolic numbers from the first to the last trials in Study 4.

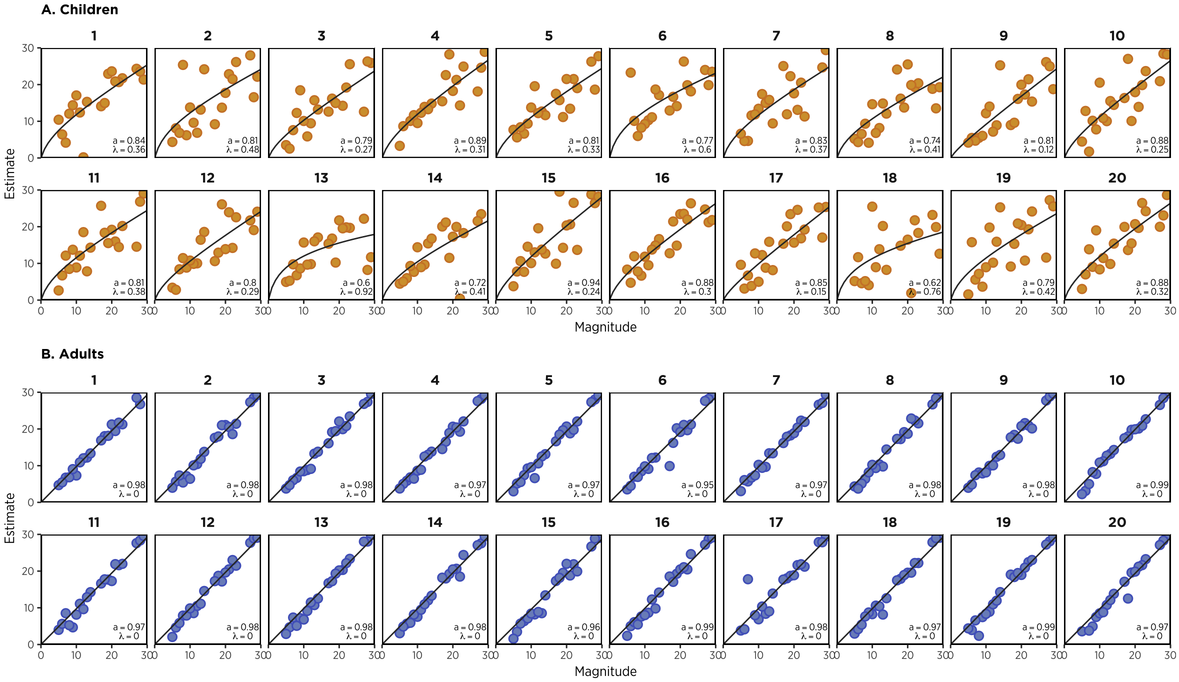

Fig. A6. Children's (A) and adults' (B) median estimates of symbolic numbers from the first to the last trials in Study 4.

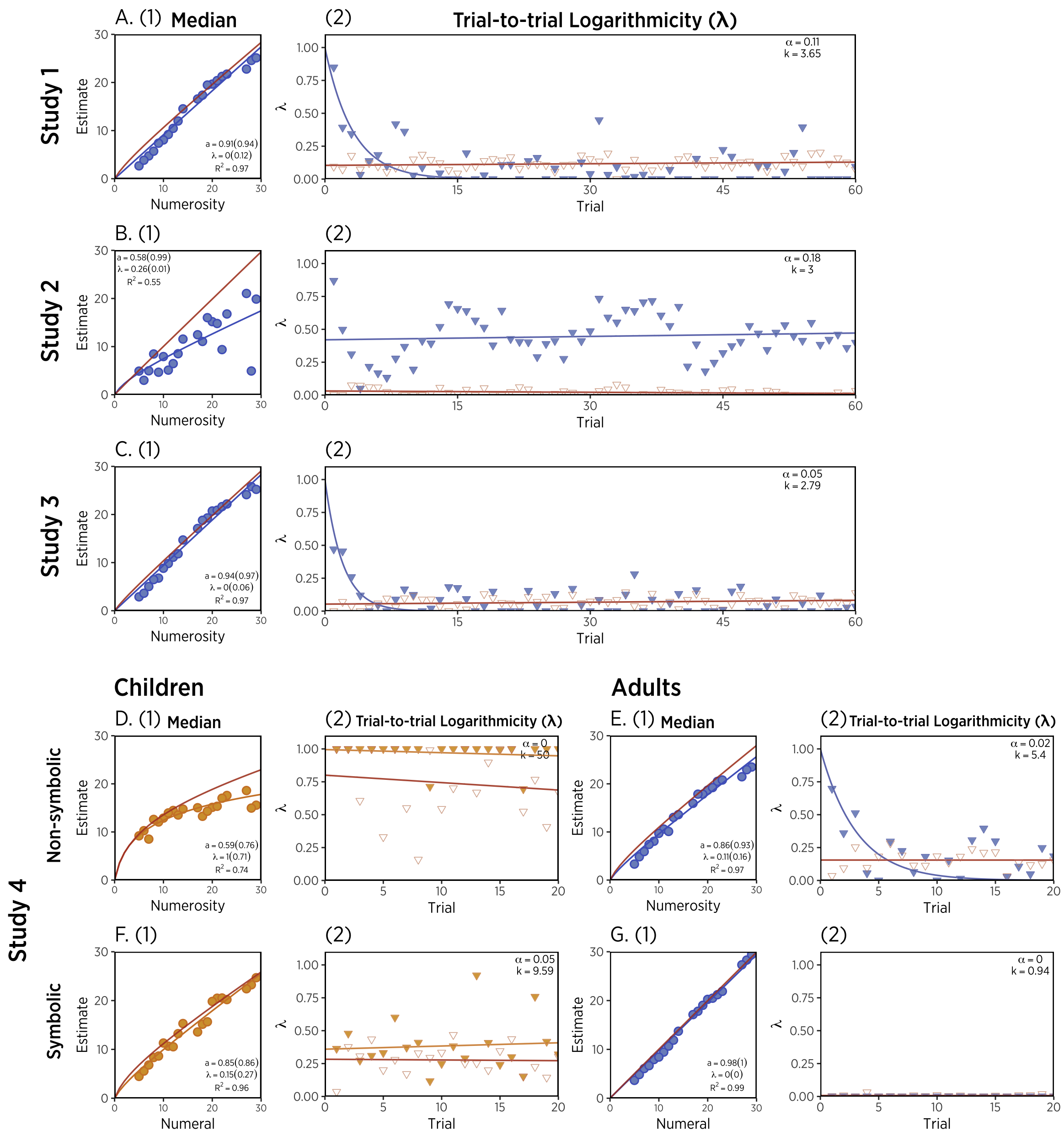

Fig. A7. Actual data (blue for adults and yellow for children) and predictions of Cicchini et al. (2014)'s model (red) in Study 1–3 (A-C) and Study 4 (D and E for non-symbolic number and F and G for symbolic number): (1) median estimates of actual and simulated data. Parameter values of the best-fitting models are shown for actual data along with those for simulated data in parentheses. (2) Trial-to-trial logarithmicity of real and simulated estimates, with parameter values of *α* and *k* used in simulations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

References

Adachi, I. (2014). Spontaneous spatial mapping of learned sequence in chimpanzees: Evidence for a SNARC-like effect. *PLoS ONE, 9*, e90373.
Anobile, G., Cicchini, G. M., & Burr, D. C. (2012). Linear mapping of numbers onto space requires attention. *Cognition, 122*, 454–459.
Anobile, G., Stievano, P., & Burr, D. C. (2013). Visual sustained attention and numerosity sensitivity correlate with math achievement in children. *Journal of Experimental Child Psychology, 116*, 380–391.
Ashcraft, M. H., & Moore, A. M. (2012). Cognitive processes of numerical estimation in children. *Journal of Experimental Child Psychology, 111*, 246–267.
Barth, H. C., & Paladino, A. M. (2011). The development of numerical estimation: Evidence against a representational shift. *Developmental science, 14*(1), 125–135.
Berteletti, I., Lucangeli, D., Piazza, M., Dehaene, S., & Zorzi, M. (2010). Numerical estimation in preschoolers. *Developmental Psychology, 41*, 545–551.

Cantlon, J. F., Cordes, S., Libertus, M. E., & Brannon, E. M. (2009). Comment on" Log or Linear? Distinct Intuitions of the Number Scale in Western and Amazonian Indigene Cultures". *Science, 323*(5910), 38 38-38.
Cicchini, G. M., Anobile, G., & Burr, D. C. (2014). Compressive mapping of number to space reflects dynamic encoding mechanisms, not static logarithmic transform. *Proceedings of the National Academy of Sciences, 111*(21), 7867–7872.
Ciesielski, K. T., Lesnik, P. G., Savoy, R. L., Grant, E. P., & Ahlfors, S. P. (2006). Developmental neural networks in children performing a Categorical N-Back Task. *Neuroimage, 33*(3), 980–990.
Crawford, L. E., Huttenlocher, J., & Engebretson, P. H. (2000). Category effects on estimates of stimuli: Perception or reconstruction? *Psychological Science, 11*(4), 280–284.
Cross, D. V. (1973). Sequential dependencies and regression in psychophysical judgments. *Perception & Psychophysics, 14*(3), 547–552.
de Hevia, M. D., Izard, V., Coubart, A., Spelke, E. S., & Streri, A. (2014). Representations of space, time, and number in neonates. *Proceedings of the National Academy of Sciences of the United States of America, 111*, 4809–4813.
de Hevia, M. D., & Spelke, E. S. (2010). Number-space mapping in human infants. *Psychological Science, 21*, 653–660.
de Hevia, M. D., & Spelke, E. S. (2013). Not all continuous dimensions map equally: Number-brightness mapping in human infants. *PLoS ONE, 8*, e81241.
Dehaene, S. (2003). The neural basis of the Weber-Fechner law: A logarithmic mental number line. *Trends in Cognitive Sciences, 7*, 145–147.
Dehaene, S. (2007). Symbols and quantities in parietal cortex: Elements of a mathematical theory of number representation and manipulation. In P. Haggard, & Y. Rossetti (Eds.). *Attention & performance xxii. Sensori-motor foundations of higher cognition* (pp. 527–574). Cambridge, Mass: Harvard University Press.
Dehaene, S., & Changeux, J. P. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience, 5*(4), 390–407.
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.
Dotan, D., & Dehaene, S. (2013). How do we convert a number into a finger trajectory? *Cognition, 129*, 512–529.
Dotan, D., & Dehaene, S. (2016). On the origins of logarithmic number-to-position mapping. *Psychological Review, 123*(6), 637.
Drucker, C. B., & Brannon, E. M. (2014). Rhesus monkeys (Macaca mulatta) map number onto space. *Cognition, 132*, 57–67.
Fazio, L. K., Bailey, D. H., Thompson, C. A., & Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *Journal of Experimental Child Psychology, 123*, 53–72.
Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*, 307–314.
Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences, 4*(2), 59–65.
Garner, W. R. (1953). An informational analysis of absolute judgments of loudness. *Journal of Experimental Psychology, 46*(5), 373.
Geary, D. C., Hoard, M. K., Nugent, L., & Byrd-Craven, J. (2008). Development of number line representations in children with mathematical learning disability. *Developmental Neuropsychology, 33*, 277–299.
Gebuis, T., & Reynvoet, B. (2012). The interplay between nonsymbolic number and its continuous visual properties. *Journal of Experimental Psychology: General, 141*(4), 642.
Gescheider, G. A. (1988). Psychophysical scaling. *Annual Review of Psychology, 39*(1), 169–200.
Goldstone, R. L., Medin, D. L., & Halberstadt, J. (1997). Similarity in context. *Memory & Cognition, 25*(2), 237–255.
Gunderson, E. A., Ramirez, G., Beilock, S. L., & Levine, S. C. (2012). The relation between spatial skill and early number knowledge: The role of the linear number line. *Developmental Psychology, 48*, 1229–1241.
Heine, A., Thaler, V., Tamm, S., Hawelka, S., Schneider, M., Torbeyns, J., ... Jacobs, A. (2010). What the eyes already "know": Using eye movement measurement to tap into children's implicit numerical magnitude representations. *Infant and Child Development, 19*(2), 175–186.
Helson, H. (1964). *Adaptation-level theory*. New York: Harper & Row.
Hollands, J. G., & Dyre, B. (2000). Bias in proportion judgments: The cyclical power model. *Psychological Review, 107*, 500–524.
Hollingworth, H. L. (1910). The central tendency of judgment. *The Journal of Philosophy, Psychology and Scientific Methods, 7*(17), 461–469.
Holyoak, K. J. (1978). Comparative judgments with numerical reference points. *Cognitive Psychology, 10*(2), 203–243.
Hurst, M., Leigh Monahan, K., Heller, E., & Cordes, S. (2014). 123s and ABCs: Developmental shifts in logarithmic-to-linear responding reflect fluency with sequence values. *Developmental Science, 17*, 892–904.
Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review, 98*(3), 352.
Jesteadt, W., Luce, R. D., & Green, D. M. (1977). Sequential effects in judgments of loudness. *Journal of Experimental Psychology: Human Perception and Performance, 3*(1), 92.
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.
Kaswan, J., & Young, S. (1963). Stimulus exposure time, brightness, and spatial factors as determinants of visual perception. *Journal of Experimental Psychology, 65*, 113–123.
Kim, D., & Opfer, E. J. (2017). A unified framework for bounded and unbounded numerical estimation. *Developmental Psychology, 53*(6), 1088.
Krueger, L. E. (1982). Single judgments of numerosity. *Perception & Psychophysics, 31*, 175–182.
Kwon, H., Reiss, A. L., & Menon, V. (2002). Neural basis of protracted developmental changes in visuo-spatial working memory. *Proceedings of the National Academy of Sciences, 99*(20), 13336–13341.
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.
Laski, E. V., & Yu, Q. (2014). Number line estimation and mental addition: Examining the potential roles of language and education. *Journal of Experimental Child Psychology, 117*, 29–44.
Lourenco, S. F., & Longo, M. R. (2010). General magnitude representation in human infants. *Psychological Science, 21*, 873–881.
McCrink, K., & Opfer, J. E. (2014). Development of spatial-numerical associations. *Current Directions in Psychological science, 23*, 439–445.
Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100*(2), 254.
Mori, S., & Ward, L. M. (1995). Pure feedback effects in absolute identification. *Perception & Psychophysics, 57*(7), 1065–1079.
Nieder, A. (2005). Counting on neurons: The neurobiology of numerical competence. *Nature Reviews Neuroscience, 6*, 177–190.
Nieder, A., & Dehaene, S. (2009). Representation of number in the brain. *Annual Review of Neuroscience, 32*, 185–208.
Nieder, A., & Merten, K. (2007). A labeled-line code for small and large numerosities in the monkey prefrontal cortex. *Journal of Neuroscience, 27*, 5986–5993.
Nosofsky, R. M. (1988). Similarity, frequency, and category representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(1), 54.
Opfer, J. E., & Furlong, E. E. (2011). How numbers bias preschoolers' spatial search. *Journal of Cross-Cultural Psychology, 42*, 682–695.
Opfer, J. E., & Martens, M. A. (2012). Learning without representational change: Development of numerical estimation in individuals with Williams syndrome. *Developmental Science, 15*(6), 863–875.
Opfer, J. E., & Siegler, R. S. (2007). Representational change and children's numerical estimation. *Cognitive Psychology, 55*, 169–195.
Opfer, J. E., & Siegler, R. S. (2012). Development of quantitative thinking. In K. Holyoak, & R. Morrison (Eds.). *Oxford handbook of thinking and reasoning*. New York: Oxford University Press.
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.
Opfer, J. E., & Thompson, C. A. (2008). The trouble with transfer: Insights from microgenetic changes in the representation of numerical magnitude. *Child Development, 79*, 788–804.
Opfer, J. E., Thompson, C. A., & Kim, D. (2016). Free versus anchored numerical estimation: A unified approach. *Cognition, 149*, 11–17.
Parducci, A. (1965). Category judgment: A range-frequency model. *Psychological Review, 72*(6), 407.
Peters, E., & Bjalkebring, P. (2015). Multiple numeric competencies: When a number is not just a number. *Journal of Personality and Social Psychology, 108*(5), 802.
Petrov, A. A. (2008). Additive or Multiplicative Perceptual Noise? Two Equivalent Forms of the ANCHOR Model. *Journal of Social & Psychological Sciences, 1*(2).

Petrov, A. A., & Anderson, J. R. (2005). The dynamics of scaling: A memory-based anchor model of category rating and absolute identification. *Psychological Review, 112*(2), 383.
Piantadosi, S. T. (2016). A rational analysis of the approximate number system. *Psychonomic Bulletin & Review, 23*(3), 877–886.
Piantadosi, S. T., & Cantlon, J. F. (2017). True numerical cognition in the wild. *Psychological Science, 28*(4), 462–469.
Qin, J., Kim, D., & Opfer, J. E. (2017). Varieties of numerical estimation: A unified framework. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.). *Proceedings of the 39th annual meeting of the cognitive science society* (pp. 2943–2948). Austin, TX: Cognitive Science Society.
Ratcliff, R., Love, J., Thompson, C. A., & Opfer, J. E. (2012). Children are not like older adults: A diffusion model analysis of developmental changes in speeded responses. *Child Development, 83*(1), 367–381.
Rouder, J. N., & Geary, D. C. (2014). Children's cognitive representation of the mathematical number line. *Developmental Science, 17*, 525–536.
Rugani, R., Vallortigara, G., Priftis, K., & Regolin, L. (2015). Number-space mapping in the newborn chick resembles humans' mental number line. *Science, 347*, 534–536.
Sasanguie, D., De Smedt, B., Defever, E., & Reynvoet, B. (2012). Association between basic numerical abilities and mathematics achievement. *British Journal of Developmental Psychology, 30*, 344–357.
Sasanguie, D., Göbel, S. M., Moll, K., Smets, K., & Reynvoet, B. (2013). Approximate number sense, symbolic number processing, or number-space mappings: What underlies mathematics achievement? *Journal of Experimental Child Psychology, 114*, 418–431.
Schley, D. R., & Peters, E. (2014). Assessing "economic value": Symbolic-number mappings predict risky and riskless valuations. *Psychological Science, 25*, 753–761.
Sella, F., Berteletti, I., Lucangeli, D., & Zorzi, M. (2015). Varieties of quantity estimation in children. *Developmental Psychology, 51*, 758–770.
Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development, 75*, 428–444.
Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14*, 237–243.
Siegler, R. S., Thompson, C. A., & Opfer, J. E. (2009). The logarithmic-to-linear shift: One learning sequence, many tasks, many time scales. *Mind, Brain, and Education, 3*, 143–150.
Sjöberg, L., & Thorslund, C. (1979). A classificatory theory of similarity. *Psychological Research Psychologische Forschung, 40*(3), 223–247.
Slusser, E. B., Santiago, R. T., & Barth, H. C. (2013). Developmental change in numerical estimation. *Journal of Experimental Psychology: General, 142*(1), 193.
Spence, I. (1990). Visual psychophysics of simple graphical elements. *Journal of Experimental Psychology: Human Perception and Performance, 16*(4), 683.
Staddon, J. E., King, M., & Lockhead, G. R. (1980). On sequential effects in absolute judgment experiments. *Journal of Experimental Psychology: Human Perception and Performance, 6*(2), 290.
Stevens, S. S. (1975). *Psychophysics: Introduction to its perceptual, neural, and social prospects.* New York: Wiley.
Sullivan, J., & Barner, D. (2014). The development of structural analogy in number-line estimation. *Journal of Experimental Child Psychology, 128*, 171–189.
Thompson, C. A., & Opfer, J. E. (2008). Costs and benefits of representational change: Effects of context on age and sex differences in symbolic magnitude estimation. *Journal of Experimental Child Psychology, 101*(1), 20–51.
Thompson, C. A., & Opfer, J. E. (2010). How 15 hundred is like 15 cherries: Effect of progressive alignment on representational changes in numerical cognition. *Child Development, 81*, 1768–1786.
Thompson, C. A., & Opfer, J. E. (2016). Learning linear spatial-numeric associations improves accuracy of memory for numbers. *Frontiers in Psychology, 7*, 24.
Thompson, C. A., Ratcliff, R., & McKoon, G. (2016). Individual differences in the components of children's and adults' information processing for simple symbolic and non-symbolic numeric decisions. *Journal of Experimental Child Psychology, 150*, 48–71.
Thompson, C. A., & Siegler, R. S. (2010). Linear numerical-magnitude representations aid children's memory for numbers. *Psychological Science, 21*, 1274–1281.
Tversky, A. (1977). Features of similarity. *Psychological Review, 84*(4), 327.
Verguts, T., & Fias, W. (2004). Representation of number in animals and humans: A neural model. *Journal of Cognitive Neuroscience, 16*, 1493–1504.
Ward, L. M., & Lockhead, G. R. (1970). Sequential effects and memory in category judgments. *Journal of Experimental Psychology, 84*(1), 27.
White, S. L., & Szűcs, D. (2012). Representational change and strategy use in children's number line estimation during the first years of primary school. *Behavioral and Brain Functions, 8*(1), 1.