Why Children's Number-line Estimates Follow Fechner's Law

Frank Kanayet (kanayet.1@osu.edu)
Department of Psychology, 255 Psychology Building
Columbus, OH 43206 USA

John Opfer (opfer.7@osu.edu)
Department of Psychology, 245 Psychology Building
Columbus, OH 43206 USA

Abstract

Estimates of numerical magnitude in young children and Amazon indigene have been observed to follow Fechner's Law, with estimates increasing logarithmically with actual value. Two models have been proposed to account for this data. The logarithmic model depicts numeric magnitudes as scaled logarithmically with constant Gaussian variability, whereas the accumulator model depicts them as scaled linearly with increasing variability. This paper tests these models by examining number-line estimation with novel magnitudes and ranges (0-100, 0-1000, 900-1000, 900-1900). Results suggest that although both models provide good fits for estimates on 0-1000 number lines, only the fit of the logarithmic model generalizes to estimates for smaller intervals (900-1000) and larger numbers (900-1900).

Keywords: Numerical cognition; representation; mathematical modeling; conceptual development.

Introduction

Whether tracking the size of a sheep flock or traveling to the moon, humans must code their experiences numerically. Even simple tasks—like matching cardinality of sets, discriminating between sets that differ only in numerosity, or performing approximate arithmetic operations—require that subjects represent numeric value. In this paper, we test two prominent models that have been proposed to characterize early mental representations of numeric value.

Among preschoolers, human infants, and non-human animals, numeric representations typically follow Fechner's Law, with discrimination between numerosities depending on the ratio of quantities and not absolute difference (Dehaene, Dehaene-Lambertz & Cohen, 1998; Gallistel & Gelman, 2000). Among older children and adults, comparisons of Arabic numbers also follow Fechner's Law (Moyer & Landauer, 1967; Sekuler & Mierkewicz, 1977), with subjects being slower and less accurate when comparing numbers that differ in distance (e.g., 9 and 7 *vs* 9 and 5) or size (e.g., 9 and 7 *vs* 5 and 3).

Two models of numeric representations are consistent with these size and distance effects: the logarithmic model (Dehaene & Changeaux, 1993) and the accumulator model (Gibbon & Church, 1981). The logarithmic model explains size and distance effects by proposing that numerical magnitudes are represented in a logarithmically-compressed scale with constant Gaussian variability. The accumulator model proposes that numbers are represented in a linear scale with variability increasing with numeric value (i.e., with scalar variability). Both models can account for size and distance effects because in each model the amount of signal overlap between representations of any two numbers is a function of the size and distance between them.

Development of Number-line Estimation

To provide a novel test of the logarithmic and accumulator models, Siegler and Opfer (2003) asked children (2nd, 4th and 6th graders) and adults to estimate the position of numbers on a blank line flanked by two numbers (0-100 or 0-1000). Assuming a one-to-one mapping between position estimates and internal representations, the number-line task provides a straightforward test of the two models. Consistent with the logarithmic model (and Fechner's Law), estimates for the 0-1000 line of nearly all 2nd graders and roughly half of 4th graders increased logarithmically with actual number. In contrast, estimates of almost all adults and 6th graders increased linearly with actual value (and without scalar variability).

This logarithmic-to-linear shift in number-line estimation has since been replicated with children of different ages (Booth & Siegler, 2006; Opfer & Thompson, 2008), cultures (Dehaene et al., 2008; Siegler & Mu, 2008) and experimental tasks (Opfer & DeVries, 2008; Thompson & Opfer, 2008). Booth and Siegler (2006), for example, found that nearly all kindergartners' estimates increased logarithmically in the 0-100 task, whereas roughly half of 1st graders and nearly all 2nd graders' estimates increase linearly. Recently, Dehaene et al. (2008) also gave the number-line task to Munducuru participants– an Amazonian tribe that had little contact with formal education and a limited numeric vocabulary (Dehaene et al., 2006; Pica et al., 2004). Consistent with the logarithmic-to-linear shift hypothesis, Dehaene et al. found that adult Munducuru participants—like kindergartners and 1st graders in Booth and Siegler's (2006) study—mapped numbers to space logarithmically, whereas Portuguese-schooled members of the Munducuru—like adults in Siegler and Opfer's (2003) study—mapped numbers to space linearly. Further suggesting that experience, not maturation, accounted for the logarithmic-to-linear shift, studies have shown that children will switch from the logarithmic to linear estimation patterns when given feedback on a single estimate that has a large discrepancy from the linear pattern

(Opfer & Siegler, 2007; Opfer & Thompson, 2008; Thompson & Opfer, 2008).

A straightforward account of how representations of symbolic number develop is implied by previous studies of number-line estimation. In this view, early representations of symbolic number are initially logarithmically compressed, much like non-symbolic numeric representations in pigeons (Roberts, 1995), rats (Gallistel, 1990), and chimpanzees (Boysen & Bernston, 1989). With age, children in many societies experience feedback on the unreliability of logarithmic representations. Because these experiences will largely entail feedback on smaller numbers (due to their disproportionate frequency in text and speech; Dehaene & Changeux, 1993), linear representations will develop sooner for small numeric ranges than for large ones. Thus, while children eventually develop linear representations, these representations do not fully supplant logarithmic ones.

Why Do Children's Number-line Estimates Follow Fechner's Law?

The developmental account offered above has been recently challenged by two proposals defending the accumulator model. First, Cantlon et al. (2009) have recently defended the accumulator model by claiming that a linearly-scaled representation of number actually predicts logarithmic—not linear—performance on the number-line task. Within this account, participants do not report the psychological distance between the probes and anchors (which would be linear in their model), but instead report the similarity ratio between the probes and anchors (Cantlon et al., 2009). Thus, because the signal overlap between 150 and 1000 is greater than the signal overlap between 1 and 150, 150 should seem more similar to 1000 than to 1, just as on a logarithmic scale.

The second defense of the accumulator model has come from Ebersbach et al. (2008), who proposed that logarithmic performance on the number-line task can be explained by a *segmented* linear model, with the two linear segments having different slopes and appearing logarithmic. From their perspective, the difference in slopes is explained by children being more familiar with small numbers than large ones. In this way, children's performance on the 0-100 and the 0-1000 would differ because children can count in the range from 0 to 100 but are unfamiliar with numbers from 101 to 1000 and thus see them as having approximately equal value.

Present Studies

To test the models proposed by Siegler & Opfer (2003), Cantlon et al. (2009), and Ebersbach et al. (2008), we asked second graders to estimate the position of numbers on four different number lines (0-100, 0-1000, 900-1000, and 900-1900). Although each model predicts approximately a logarithmic pattern of estimates on 0-1000 number lines, the models make competing predictions for remaining number lines. An interesting aspect of our design was that it allowed us to examine estimates for the same numbers appearing in different contexts, thereby allowing us to test the effect of numeric magnitude on estimates (Table 1). On 0-1000 number lines, 900-1000 are at the end of the interval, and all models predict compression for these estimates. On the other hand, in the 900-1900 line, the same numbers are at the beginning of the interval and only the logarithmic model predicts less compression for this context. Finally, in the 900-1000 line, the large numbers are distributed in the whole range, leading the logarithmic model to predict nearly linear estimates.

What do these three models predict for each interval? Predictions of logarithmic-to-linear shift model were derived by assuming logarithmic scaling that would pass through the two endpoints of the scale. For the 0-100 task the predictions follow the function $y = x$. For the 0-1000 task the predictions follow the function $y = 144.76 * \ln(x)$. For the 900-1000 task predictions follow the function $y = 929.12 * \ln(x)$. Finally, for the 900-1900 task predictions follow the function $y = 1338.3 * \ln(x)$.

To derive predictions of Cantlon et al. (2009) model, we modeled scalar variability by assuming numerical magnitude was equal to Gaussian noise in representations. (Scalar variability allows for any noise-to-magnitude ratio, but different ratios affect only the intercepts, not overall shape. Further, our assumption is generous in that it mimics logarithmic scaling.) After defining the Gaussian noise for each number, children's estimates were predicted first by calculating overlap between the Gaussian distribution of linearly-scaled probe values and each of the two anchors (Equation 1 and Equation 2), and then by calculating the ratio of those two similarity values (Equation 3).

$$x = \frac{-\frac{(\sigma_1^2 * \mu_2) - (\sigma_2^2 * \mu_1)}{\sigma_1^2 * \sigma_2^2} \pm \sqrt{\left(\frac{(\sigma_1^2 * \mu_2) - (\sigma_2^2 * \mu_1)}{\sigma_1^2 * \sigma_2^2}\right)^2 - 4\left(\frac{\sigma_2^2 - \sigma_1^2}{2 * \sigma_1^2 * \sigma_2^2}\right)\left(\frac{(\sigma_2^2 * \mu_1^2) - (\sigma_1^2 * \mu_2^2)}{2 * \sigma_1^2 * \sigma_2^2} - \ln\left(\frac{\sigma_2}{\sigma_1^2}\right)\right)}}{2\left(\frac{\sigma_2^2 - \sigma_1^2}{2 * \sigma_1^2 * \sigma_2^2}\right)} \quad (1)$$

$$\text{Similarity} = c_1(x_1) + c_2(x_2) - c_2(x_1) + 1 - c_1(x_2) \quad (2)$$

Where c_1 is the cumulative density function for the distribution of the lower number, c_2 is the cumulative density function for the larger number, x_1 the first intersection between the two distributions, and x_2 the second intersection between the distributions. Children's actual estimates were predicted by *y*.

$$Y = \frac{\text{Similarity}_1}{\text{Similarity}_1 + \text{Similarity}_2} \quad (3)$$

Where Similarity_1 is the degree of overlap between the distribution of the probe and the lower anchor and Similarity_2 is the degree of overlap between the probe and the higher anchor.

To derive predictions of Ebersbach's model, we used ordinary least squares to find the best-fitting segmental linear function for a data series (0, 1000) defined by $y = 144.76 * \ln(x)$. Then, using the four parameter-values from this function, we calculated predicted estimates for remaining tasks. The theoretical justification for this strategy is that the inflection point, x_0, is thought to reflect children's prior experience with numbers, rather than task-specific factors. Additionally, our best-fitting segmental linear function obtained parameter values that were very close to those observed by Ebersbach et al (2008).

The key test for the models was how accurately they predicted estimates for large numbers with small intervals (900-1000) and large intervals (900-1900) (for predictions, see Figs. 1 – 4). Specifically, both the linear model with scalar variability (Cantlon et al., 2008) and the segmented-linear model (Ebersbach et al., 2008) predict compression among large numeric values, regardless of context. Within the Cantlon et al. (2009) account, compression stems from the large overlap in noise among large numbers. Thus, Cantlon et al.'s (2009) model predicts that estimates will increase logarithmically on 0-100 and 0-1000 number lines, whereas it predicts estimates will cluster around 950 on 900-1000 number lines and will cluster between 1300 to 1500 on 900-1900 number lines. Within Ebersbach et al.'s (2008) account, compression stems from unfamiliarity of numbers beyond a certain range. For numbers that fall outside of children's familiarity (x_0), the model predicts estimates will cluster between 990-1030 on 900-1000 number lines and cluster between 990-1400 on 900-1900 number lines. In contrast, logarithmic scaling depicts the magnitude of compression as a function of interval size.

Method

Participants

Participants included 17 American 2nd graders (*N* = 17; *M* = 8.28, *SD* = 0.31; 9 females, 8 males). One child was excluded for inattention and 5 children excluded for providing adult-like estimates on 0-100 and 0-1000 line. In this way, we could focus exclusively on the fit of non-linear models (i.e., segmental linear, logarithmic, and accumulator).

Design and Procedure

Each child was presented with all four types of number-lines (0-100, 0-1000, 900-1000, and 900-1900), with type of number line counterbalanced using a Latin-square design. Problems were presented on a computer screen; each line had a width of 255 pixels.

At the beginning of each trial a fixation was presented 192 pixels over the line (half point between the top of the screen and the number line) for one second. Afterwards, the number probe appeared at fixation, and children had to mouse click to indicate the position of the number. Children were instructed to answer with both speed and accuracy in mind. For each task, participants estimated positions of 20 numbers, one per line (see Table 1). Numbers were chosen to sample over the whole range, to minimize effects of specific knowledge (e.g., that 50 is half of 100), and to over-sample at the low end of the range (where models most strongly diverge in predictions).

Table 1: Numbers estimated on number lines (0-100, 0-1000, 900-1000, 900-1900)

0-100	0-1000	900-1000	900-1900
1	1	901	901
2	5	902	905
4	10	904	910
5	26	905	926
8	47	908	947
10	68	910	968
12	90	912	990
13	130	913	1030
15	150	915	1050
17	260	917	1160
26	470	926	1370
32	680	932	1580
37	700	937	1600
47	830	947	1730
58	905	958	1805
68	910	968	1810
70	926	970	1826
83	947	983	1847
90	968	990	1868
94	990	994	1890

Results and Discussion

Model Comparison

To assess the general function of children's numerical estimates, we first examined median estimates of participants for each of the four number lines, and we compared the fits of the linear and logarithmic regression models to these estimates. Consistent with the linear-to-logarithmic shift hypothesis, estimates on 0-100 number lines were better fit by the linear regression function than by the logarithmic (lin R^2 = .98; log R^2 = .80), whereas the logarithmic function provided the better fitting function on 0-1000 number lines (lin R^2 = .83; log R^2 = .91), 900-1000 number lines (lin R^2 = .92; log R^2 = .93), and 900-1900 number lines (lin R^2 = .87; log R^2 = .90).

The observed pattern of estimates is not consistent with the two models inspired by the accumulator model. According to Cantlon et al.'s (2009) version, a logarithmic model would provide the best fit across all four tasks, yet estimates on the 0-100 task were best fit by the linear function. Additionally, Ebersbach et al.'s (2008) version predicts that a linear function would provide the best fit for estimates on 900-1000 and 900-1900 lines, but the logarithmic model provided a better fit for the 900-1000 and 900 – 1900 number lines.

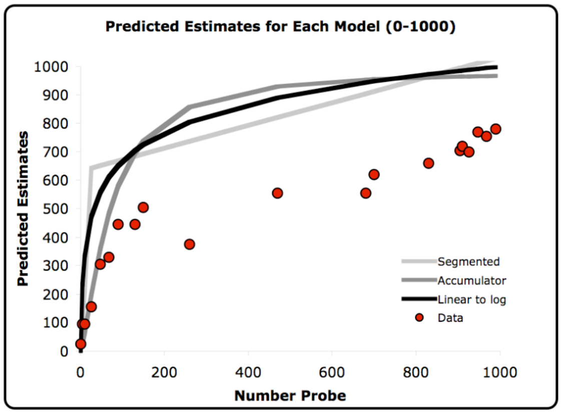

Figure 1: Predictions and median estimates for the 0-1000 task

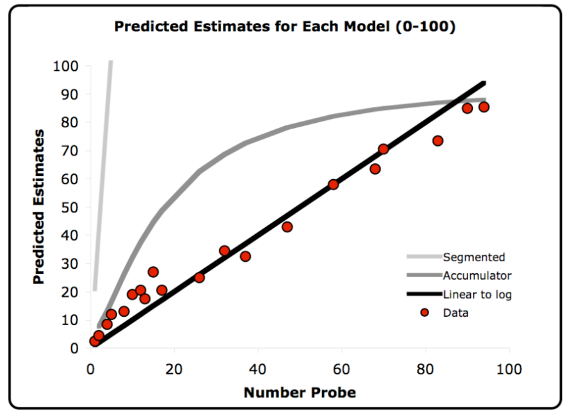

Figure 2: Predictions and median estimates for the 0-100 task

We next examined how well each model predicted the specific values of children's estimates. To measure this, we first calculated the predicted estimate for each numeric value, and we then measured the mean absolute error (*MAE*) of the predicted estimate to children's median estimate.

We first evaluated performance on the 0-1000 number line (Fig 1), where we expected the three models to yield very similar predictions. For the Siegler & Opfer (2003) model, we used the ideal logarithmic function,,$\hat{Y} = 144.6 * \text{Ln}(x)$, to predict children's median estimates. For Cantlon et al.'s (2009) model, we used Equations 1 -3. For Ebersbach's model, parameters for the segmented-linear model were obtained by finding the least squares best fitting function to the logarithmic series with the additional constraint that the first intercept should be equal to 0 ($a_1 = 0$; $b_1 = 20.89$; $b_2 = .4$; $x_0 = 30.87$).

As illustrated in Figure 1, the three models predicted children's estimates on the 0-1000 line equally well, though all three models tend to over-estimate children's estimates. Specifically, we found a *MAE* = 257.13 (*SD* = 113.72) for the segmented-linear model; a *MAE* = 222.17 (*SD* = 125.14) for the accumulator model; and a *MAE* =262.29 (*SD* = 86.46) for the logarithmic model (See Figure 2). A one-way (model: logarithmic, accumulator, segmented-linear) ANOVA revealed that the main effect of model was not statistically significant ($F(2, 56) = 0.77$, $p = .468$).

One of the most important criteria for model selection, however, is the ability to generalize to novel tasks (Pitt & Myung, 2002): good models achieve a good fit because they fit signal, not because they overfit noise. Thus, the strongest test for model selection is looking at how well the models generalize to the 0-100, 900-1000 and 900-1900 tasks.

For the 0-100 task (Fig. 2), a one-way (model) ANOVA showed a significant effect of model on *MAE*, $F(2, 56) =18.85$, $p < .001$. A multiple comparisons post-hoc analysis revealed that the linear model (*MAE* = 4.9, *SD* = 3.3) predicted children's estimates significantly better than the accumulator model (*MAE* = 18.37, *SD* = 12.33), which in turn outperformed the segmented-linear model (*MAE* = 383.44, *SD* = 234.39). The superior fit of the linear model y = x) to the accumulator model is not surprising given previous results, an observations also made by Dehaene et al. (2009). The poor fit of the segmented linear model to the 0-100 task is more interesting: clearly the segmented linear model achieves a good fit on the 0-1000 task by overfitting the data.

Children's estimates on 900-1000 and 900-1900 number lines had not been observed previously, and they provided an excellent opportunity to test the generalizability of the three models to new data. Additionally, they provided an opportunity to examine estimation on number lines with the same interval size as in previous studies but with greater magnitude. We predicted that the logarithmic model would perform better in these ranges because it proposes that numbers are represented in a logarithmic scale, regardless of the magnitude. In contrast, both versions of the accumulator model predict that when numbers are large enough there is so much noise in the representation that numbers become virtually indistinguishable.

For the 900-1000 task (Fig. 3), there was again a significant effect of model on *MAE* scores ($F(2, 56) = 182.09$ $p < .001$). Post-hocs revealed that the logarithmic model (*MAE* = 8.21, *SD* = 8.96) predicted children's estimates significantly better than the accumulator model (*MAE* = 23.3, *SD* = 14.31), and in turn the accumulator model predicted children's estimates significantly better than the segmented-linear model (*MAE* = 75.08, *SD* = 11.08).

For the 900-1900 task (Fig. 4), models also differed in how well they predicted children's estimates, $F(2, 56) = 94.34$ $p < .001$. Post-hoc analysis revealed that the logarithmic model (*MAE* =150.7, *SD* = 72.17) and the accumulator model (*MAE* =172.7, *SD* = 84.5) predicted significantly better the performance in the task than the segmented-linear model (*MAE* = 232.75, *SD* = 119.23), though there was a nominal superiority for the logarithmic model.

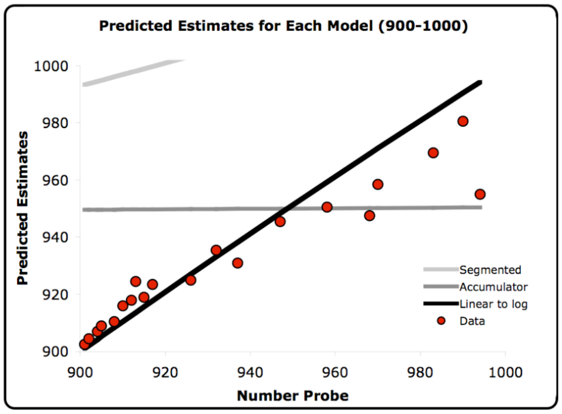

Figure 3: Predictions and median estimates for the 900-1000 task

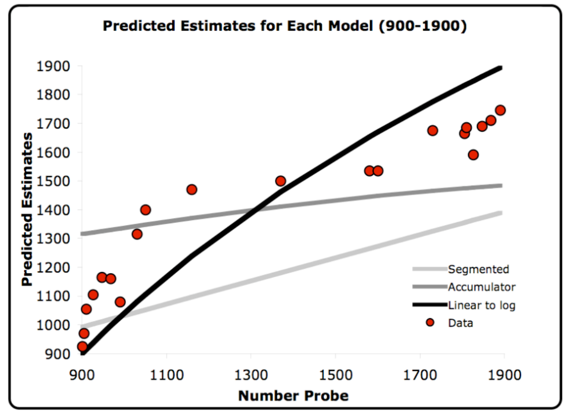

Figure 4: Predictions and median estimates for the 900-1900 task

Overall, the results of the model comparisons favor the developmental account offered by the linear-to-logarithmic shift hypothesis, and results undermine the two recent defenses of the accumulator model offered by Canton et al. (2009) and Ebersbach et al. (2008). In the 0-100 task, this is the only model that clearly predicts a linear performance and the results showed that children do perform linearly in this range of numbers. As noted before, the segmented-linear model could predict this performance by estimating the four free parameters of the model for the range. However, this would cost them in the performance in the 0-1000 line. We argue that because the rationale behind this model is that the second slope corresponds to the degree of familiarity with each number, the parameter values found in one task should be used for the others, and thus, is not possible for this model to avoid the tradeoffs presented by this combination of tasks. This means that, in general the segmented-linear model would be the best performer for the particular range used to estimate the parameters, but would be the worse in the other three intervals. Finally, the accumulator model commits to a logarithmic performance in the number-line task and that immediately disqualifies it for the 0-100 range.

The linear-to-logarithmic model also provided more accurate predictions of children's estimates in the 900-1000 and 900-1900 ranges. It is worth highlighting that the model predicted the performance in the 900-1000 line remarkably well. In general, although the overall mean absolute errors are similar between the logarithmic model and the accumulator model, for these two tasks, a visual inspection of the performance shows that the logarithmic model captures the pattern of the data much better. In the case of the accumulator model, for large numbers, the amount of noise is so large that makes the number probes virtually indistinguishable, and although their proposed process of estimation takes the anchors into account, it is not enough to fit the data well. In these ranges, the performance of the segmented-linear model misses completely because the model's assumptions relating the compression with lack of familiarity and proposing a linear slope *ad infinitum* leaves this model without the ability to generalize to novel intervals with larger numbers.

Finally, we found no significant differences between the three models in the 0-1000 range. Our findings confirm the claims made by the three models and replicate previous results in estimation tasks in this range, for this particular age. This finding is also important, because it shows that the methods used in this paper to calculate the predictions of the three models find results that are consistent with the claims of the three camps and with the general understanding regarding the difficulty to disentangle the predictions of the models with behavioral data.

Effect of Interval Size and Magnitude on Numerical Representations

Our experimental design also allowed us to address a novel empirical question: What is the effect of interval size and numeric magnitude on children's estimates? To address this question, we next examined estimates for seven numbers between 900 and 1000 that were presented across three different tasks. This analysis provides an opportunity to test a central prediction of the linear-to-logarithmic hypothesis, which scales logarithmic representations to specific tasks. Specifically, the model predicts that the slope of estimates for numbers between 900 and 1000 should be steeper for the 900-1900 task than for the 0-1000 task.

Consistent with this prediction, a paired-sample *t*-test showed a significance difference between the slope of the 0-1000 line (M = .9, SD = 1.6) and the 900-1900 line (M = 3.27, SD = 1.47) $t(10) = 3.18$, $p < .001$. As noted before, the segmented-linear model does not predict a change in the slope after x_0 yet results show that slopes differ by task. In this sense, both the Siegler and Opfer (2003) and Cantlon et al. (2008) models predicted this overall pattern. Further, the Siegler & Opfer (2003) model more accurately predicted the magnitude of this slope ($b_{(0\text{-}1000)}$ =.15 vs. $b_{(900\text{-}1900)}$ = 1.42) than did the Cantlon et al. (2008) model ($b_{(0\text{-}1000)}$ = .03 vs. $b_{(900\text{-}1900)}$ = .32).

Our fine-grained analysis of estimates complements previous findings of context effects by Siegler and Opfer (2003). Specifically, Siegler and Opfer (2003) demonstrated that estimates of 0 to 100 on 0-100 number lines increased linearly, whereas estimates of the same values on 0-1000 number lines increased logarithmically, suggesting that the magnitude of the interval affects whether children access linear or logarithmic representations. The current experiment replicated the interval effect, but it also discovered that estimates for large numbers on number lines of small intervals (e.g., 900-1000 number lines) do increase logarithmically. This finding is important because it suggests that children's representations of large numbers improve over time, not just their knowledge about the demands of small versus large numeric intervals.

In summary, the best explanation for why children's number-line estimates follow Fechner's Law is not because variability in their representations increases proportionally with numbers, nor because children can only discriminate a small set of numbers for which they are familiar. Rather, children's early representations of numeric value increase logarithmically with actual value.

Acknowledgments

We thank Simon Dennis for his help in the development of Equations 1 - 3.

References

Booth, J. & Siegler, R. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 42*(1), 189-201.

Boysen, S. T. & Berntson, G. G. (1989). Numerical Competence in a Chimpanzee (*Pan troglodytes*). *Journal of Comparative Psychology, 103*(1), 23-31.

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)*,* 38b.

Dehaene, S., & Changeux, J. P. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience, 5*, 390–407.

Dehaene, S., Dehaene-Lambertz, G. & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neuroscience, 21*(8), 355-361.

Dehaene, S., Izard, V., Pica, P. & Spelke, E. (2006). *Science, 311*(5759), 381-384.

Dehaene, S., Izard, V., Pica, P. & Spelke, E. (2009). Response to Comment on "log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures". *Science, 323*(5910)*,* 38c.

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*(5880)*,* 1217-1220.

Ebersbach, M., Luwel, K., Frick, A., Onghena, P. & Verschaffel, L. (2008). The relationship between the shape of the mental number line and familiarity with numbers in 5- to 9-year old children: Evidence for a segmented linear model. *Journal of Experimental Child Psychology, 99*(1), 1-17.

Gallistel, C. R. (1990). *The Organization of Learning*. Cambridge, MA: MIT Press.

Gallistel, C. R. & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences, 4*(2), 59-65.

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

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature*, *215*, 1519-1520.

Opfer, J. E., & DeVries, J. M. (2008). Representational change and magnitude estimation: why young children can make more accurate salary comparisons than adults. *Cognition, 108*(3), 843-849.

Opfer, J. E. & Siegler, R. S. (2007). Representational change and children's numerical estimation. *Cognitive Psychology, 55*(3), 169-195.

Opfer, J. E., & Thompson, C. A. (2008). The trouble with transfer: Insights from microgenetic changes in the representation of numerical magnitude. *Child Development*, *79*, 790 -806.

Pica, P., Lemer, C., Izard, W. & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*(5695), 499-503.

Pitt, M. A. & Myung, I J. (2002). When a good fit can be bad. *Trends in Cognitive Sciences, 6*(10), 421-425.

Roberts, W. A. (1995). Simultaneous numerical and temporal processing in the pigeon. *Current Directions in Psychological Science, 4*, 47-51.

Sekuler, R., & Mierkiewicz, D. (1977). Children's judgments of numerical inequality. C*hild Development, 48*, 630–633.

Siegler, R. & Mu, Y. (2008). Chinese children excel on novel mathematics problems even before elementary school. *Psychological Science, 19*(8), 759-763.

Siegler, R. S. & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14*(3), 237-243.

Thompson, C. A., & Opfer, J. E. (2008). Costs and benefits of representational change: Effect of context on age and sex differences in magnitude estimation. *Journal of Experimental Child Psychology, 101*, 20 – 51.