Straightening Up: Number Line Estimates Shift from Log to Linear with Additional Information

Percival G. Matthews (pmatthew@nd.edu)

Dana Chesney (Dana.Chesney.3@nd.edu)
University of Notre Dame, Department of Psychology,
118 Haggar Hall, Notre Dame, IN 46556

Abstract

It has been suggested that a developmental log-to-linear shift in children's performance on number line estimation tasks is diagnostic of their underlying representations of numerical magnitude (Siegler & Opfer, 2003). However, in the study presented herein, we were able to induce a similar log-to-linear shift on number line estimation tasks among adults by manipulating their familiarity with the numbers we used as stimuli. We offer this evidence as an existence proof that differences in performance on number line estimation tasks may not necessarily be indicative of fundamental differences in the formats of people's underlying numerical magnitude representations. Rather, they may be diagnostic of differences in people's understandings of what magnitudes are represented by symbolic numbers.

Keywords: Number line, Number, Numeracy, Estimation Approximate number system

Numerical estimation is an important part of mathematical cognition for both children and adults. Indeed, numerical estimation seems to play a central role in a wide range of mathematical activities (see Siegler & Booth, 2005 for a review). In recognition of its importance, Siegler and colleagues have focused on the ability to estimate numerical magnitudes as a key indicator of number sense in developmental studies (Siegler & Booth, 2004; Siegler & Opfer, 2003)

Perhaps the most well known of these numerical estimation tasks is the number line estimation task. This task requires that participants estimate the location of a number on a line with numerical anchors at each end. This task not only involves magnitude estimation, but also the ability to translate between symbolic numbers and mental magnitudes. Siegler and colleagues have helped illuminate a typical developmental trend observed on such number line estimation tasks: young children show a tendency to compress numbers logarithmically, whereas adults do not (e.g., Siegler & Booth, 2004; Siegler & Opfer, 2003). That is, when asked to mark the position where a number n is on the line, children typically place the mark at a spot approximated by $\log(n)$ instead of n. With development, these estimates become more linear (Siegler & Opfer, 2003). Importantly, the linearity of children's magnitude estimates on these tasks is correlated with a wide range of numeracy measures, including counting ability, number naming, digit magnitude comparison, and achievement test scores (e.g., Booth, 2005; Ramani & Siegler, 2008; Siegler & Booth, 2004).

According to the log-to-linear shift hypothesis, the logarithmic compression found in children is diagnostic of their underlying representations of numerical magnitude. This view suggests that children's performance is logarithmic because their underlying mental representations of numerical magnitudes are logarithmically compressed (Dehaene, 1997; Siegler, 2009). Siegler and colleagues further theorize that experience and schooling lead to the development of a linear representation of numerical magnitude. More specifically, they propose that though the logarithmic and linear representations continue to coexist, individuals can learn to invoke the linear representation when it is appropriate (Siegler & Opfer, 2003).

Three pieces of evidence appear to support this hypothesis. First, children's estimates become more linear with schooling and experience. Second, the linearity of children's estimates predicts a wide range of numeracy measures. These include counting ability, number naming, digit magnitude comparison, and achievement test scores, perhaps indicating that the hypothesized linear ruler is widely applied once it is developed (Siegler & Booth, 2004). Finally,, children's estimates are more linear within number ranges that are more familiar to them but remain logarithmic on larger, more unfamiliar scales (Siegler & Opfer, 2003). This in particular has been taken as evidence that the logarithmic and linear rulers coexist.

The log-to-linear shift hypothesis, however, is not without controversy. A recent critique put forth by Barth and Paladino (2010) raised questions about the interpretation of children's apparently logarithmic performance. The critique argued that the shift from logarithmic to linear performance on number line tasks is not diagnostic of some basic change in children's underlying representation of number, but is instead due to knowledge constraints interfering with the default method for completing the task. Barth and Paladino argue that number line estimation cannot properly be seen as a pure numerical estimation task. Rather, such placement tasks are actually a form of proportion judgment task – a task in which ratio between items must be evaluated. Indeed, previous literature in psychophysics has shown that estimation tasks that combine two measures in a complementary fashion such that they sum to a fixed total should be characterized as proportion judgment tasks (e.g., Hollands & Dyre, 2000; Spence, 1990; Stevens, 1957). Thus, because estimating a number's place on a number line involves both the estimate of that numbers' placement relative to the zero anchor point and of its complement's placement relative to the rightmost anchor, the task is essentially a proportion judgment. For example, when placing 25 on a 0-100 line, it should be 25 units away from 0, and 75 units away from 100, and should therefore be placed 25/(25+75), or one fourth of the total length of the line away from 0.

Proportion judgment tasks tend to yield linear relationships between the actual proportion of the stimulus presented and the judged proportion, even when using stimuli for which pure magnitude estimation tasks yield a compressive relationship between actual and perceived stimulus intensity (see Hollands & Dyre, 2000; Spence, 1990). This linear performance results because the underlying compressive function is mapped to fractionated distances according to a cyclical power model, which approximates linearity because of the reference points that are perceived linearly (Spence, 1990). Importantly, this model predicts linear performance even given a compressive underlying representation of number.

These findings raise questions regarding young children's apparently logarithmic performance on the number line task. Why, given a proportion judgment task, do children's representations appear to be nonlinear in the first place, when psychophysics – even given a logarithmic underlying representation – might predict otherwise? One possible answer is that certain assumptions of the psychophysics proportion judgment model may be violated when young children perform number line estimation tasks, impeding their use of the default comparison procedure for performing the tasks.

To judge a proportion, one must know the approximate magnitude of the whole (the rightmost anchor). Indeed, the proportion judgment model assumes that participants have access to the magnitudes at both ends of the line. Although this assumption is logical when perceptual continua are used to indicate the anchors at each end of the line (e.g., bar length on a bar graph, see Spence, 1990), this is not necessarily the case with tasks that require young children to understand the magnitudes of symbolic numbers. Essentially, people who do not have a correct understanding of the values represented by both the high and low anchor points lack the knowledge needed to fully render number line estimation tasks as proportion judgment tasks. In line with this argument, Barth and Paladino (2010) proposed a modification of the proportion judgment model positing that unknowledgeable children rescale the uppermost anchor point relative to some idiosyncratic default high end numbers. This model fit children's performance data as well as the logarithmic model favored by Siegler and colleagues (but see Barth & Paladino, 2010 for a discussion of how their model may better predict error patterns for estimates). Ebersbach, Luwel, Frick, Onghena, et al. (2008) similarly proposed that children's apparently logarithmic performance resulted from knowledge of the limits of their understanding of what numbers represent, but rather suggest that their performance could best be modeled by a two stage linear function with an overly steep slope up to the point of correct understanding, and an overly shallow slope thereafter, reflecting a clustering of unmapped values towards the end of the line, while known value occupy the bulk of the available space.

Research on children's abilities to identify symbolic numbers by name provides at least some tangential support for the knowledge constraint hypothesis. For example, young children often cannot consistently name symbolic numbers above twenty, even when they can recite those numbers as part of the count sequence (Wright, 1994; see also Clarke & Shinn, 2004). One might question whether estimates based on any unrecognizable number should have a one-to-one mapping to any particular numerical magnitude, particularly given Siegler and Opfer's (2003) finding that children's estimates are linear within familiar number ranges but look logarithmic when ranges expand to include unfamiliar numbers. This apparently logarithmic performance may be an artifact of unfamiliarity rather than the result of using two different mental representations.

The Current Study

The current experiment uses adult data to raise questions about whether performance on number line estimation tasks is diagnostic of people's underlying representations of number. It is important to note that the consequences of failure to understand the values represented by high anchor points need not be limited to children. Indeed, if Barth and Paladino (2010) are correct, and logarithmic performance is an artifact of a lack of numerical understanding, then even highly numerate adults should show logarithmic performance on such tasks under sufficiently confusing conditions. There is little to suggest that adults should rely on a logarithmic ruler in the number ranges that we presented in this task. Therefore, if logarithmic performance were found in adults, it would be natural to conclude that it is an artifact of the knowledge constraints imposed on the task, rather than a product of adults' underlying representations of number. This would serve as an existence proof that logarithmic compression on number line estimation tasks can result from knowledge constraints (i.e., a lack of numerical understanding of anchor values) rather than from use of logarithmic ruler for underlying representations of numerical magnitude (compare with Ebersbach et. al., 2008; Barth & Paladino, 2010).

We investigated whether or not imposing knowledge constraints could elicit "logarithmic" performance from adults on a number line task. We presented adults with several number line estimation tasks, some of which were designed to encourage participants to hold mistaken assumptions about the magnitude of the high-end anchor of the line. We hypothesized that performance up to the erroneously assumed anchor values would be approximately linear. We further hypothesized that participants would show some confusion when presented with stimuli that violated their expectations by exceeding the assumed high anchor value. We also expected that participants would lump these unexpectedly high value numbers together in

Table 1: Stimulus description for each estimation set.

Stimulus Set	Notation		Stimulus Values	Used in Condition
	Anchor	Stimulus		
Decimal 0-100	Decimal	Decimal	2 3 7 12 16 23 29 40 58 72 82	2
Decimal 0-1000	Decimal	Decimal	2 4 6 18 25 72 157 233 395 582 820	2
Decimal 16-32k	Decimal	Decimal	16032 16064 16096 16288 16400 17136 18480 19680 22240 25195 28960	1 & 2
Decimal 0-32k	Decimal	Decimal	64 128 192 576 800 2272 4960 7360 12480 18390 25920	1 & 2
All Exponential	Exponential	Exponential	$.002 \times 10^{4.5}$ $.004 \times 10^{4.5}$ $.006 \times 10^{4.5}$ $.018 \times 10^{4.5}$ $.025 \times 10^{4.5}$ $.072 \times 10^{4.5}$ $.157 \times 10^{4.5}$ $.233 \times 10^{4.5}$ $.395 \times 10^{4.5}$ $.582 \times 10^{4.5}$ $.820 \times 10^{4.5}$	1 & 2
Decimal Stimuli – Exponential Anchors	Exponential	Decimal	Same as *Decimal 0-32k*	1
Decimal Stimuli – Exponential Anchors Calibrated	Exponential	Decimal	Same as *Decimal 0-32k*	1
Exponential Stimuli – Decimal Anchors	Decimal	Exponential	Same as *All Exponential*	2
Exponential Stimuli – Decimal Anchors Calibrated	Decimal	Exponential	Same as *All Exponential*	2

a compressed space toward the uppermost anchor, such that performance on stimuli past the assumed high-end anchor value would yield a linear slope near zero. Finally, we hypothesized that the positively sloped section (below the assumed high anchor value) and the flatter section section (above the assumed high anchor value) together would yield a set that was well fit by a logarithmic line.

Method

This study investigated how incomplete knowledge about the magnitude of numerical anchors affect adults' performance in number line estimation tasks. We used different notational systems (i.e., standard decimal notation and exponential notation using fractional powers) to create confusion about the relative values of high endpoints anchors and to-be-placed stimuli.

Participants

Participants were 67 undergraduate students from the University of Notre Dame, participating for course credit.

Materials and Design

Participants completed the experiment individually, with all training and testing stimuli presented on iMac 5.1 computers using Superlab 4 software (Cedrus Corporation, 2007). Each problem involved a 14.7-cm long line with anchor values printed below the line at the right and the left. The numbers to be estimated appeared approximately 1.5 cm above the center of the line. Participants were asked to place a cursor at the appropriate point on the number line and to indicate their answers via mouse click. Participants were given up to 15 seconds to answer on each trial.

Similar distributions of numbers to be estimated were generated for each set of lines. Anchors and stimuli were presented in either decimal notation (e.g., 192, 576) or in exponential notation (e.g., $.006 \times 10^{4.5}$, $.018 \times 10^{4.5}$). The left end of each line was labeled "0", and the right end varied according to number set as described below. The numbers to be estimated for the 0-100 line were adapted from Barth and Paladino (2010). The numbers to be estimated for the 0-1000 line were adapted from Siegler & Opfer (2003). Stimuli for all other sets were generated by multiplying the 0-1000 stimuli by the appropriate constants so that the distributions of numbers to be estimated remained identical across scales (See Table 1). Each stimulus within each set was presented twice.

With our key manipulation, we sought to create a situation for adults that would parallel a situation in which a child might not have knowledge of the upper anchor for the number line estimation task. Pilot studies led us to settle upon the upper anchor of $.999 \times 10^{4.5}$. When confronted with this exponential notation, none of our pilot subjects correctly determined that its value was equivalent to 31,623. Rather, most assumed that it was roughly equal to 10,000. We expected that participants would perform linearly for stimuli up to the assumed anchor value (typically 10,000), but that stimuli that exceeded this value would be compressed into a small space at the right end of the line

In all, there were ten different sets of number line tasks:

Controls

- *Decimal 0-1000.* These lines presented both anchors and stimuli in decimal notation, with anchors at 0 and 1000.

Table 2: Least-squares fit information for different estimation sets.

Stimulus Set	Linear R^2	Slope	Log R^2	Best Fit
Decimal 0-100	1.00	1.04	0.76	Linear $t(10) = -6.95$, $p < .01$
Decimal 0-1000	0.99	0.98	0.63	Linear $t(10) = -4.57$, $p < .01$
Decimal 16-32k	0.99	0.99	0.74	Linear $t(10) = -8.90$, $p < .01$
Decimal 0-32k	1.00	0.90	0.71	Linear $t(10) = -4.93$, $p < .01$
All Exponential	0.99	0.95	0.63	Linear $t(10) = -4.20$, $p < .01$
Decimal Stimuli - Exponential Anchors	0.85	1.21	0.91	No Significant Difference $t(10) = .63$, $p = .55$
Decimal Stimuli – Exponential Anchors Calibrated	0.99	0.92	0.69	Linear $t(10) = -4.74$, $p < .01$
Exponential Stimuli – Decimal Anchors	0.93	0.34	0.84	No Significant Difference $t(10) = -.96$, $p = .36$
Exponential Stimuli – Decimal Anchors Calibrated	0.99	0.90	0.64	Linear $t(10) = -4.18$, $p < .01$

- *Decimal 0-100*. These lines presented both anchors and stimuli in decimal notation, with anchors at 0 and 100.
- *Decimal 0-32k*. These lines presented both anchors and stimuli in decimal notation, with anchors at 0 and 31,623.
- *Decimal 16-32k*. These lines presented both anchors and stimuli in decimal notation, with anchors at 16,000 and 31,623.
- *All Exponential*. These lines presented both anchors and stimuli in exponential notation, with anchors at 0 and $.999\text{x}10^{4.5}$

Incompatible Notation

- *Decimal Stimuli - Exponential Anchors.* These lines presented anchors in exponential notation and stimuli in decimal notation. Anchors were at 0 and $.999\text{x}10^{4.5}$.
- *Decimal Stimuli – Exponential Anchors Calibrated.* Before beginning trials within this set, participants were shown a single slide with the proper location of 16,000 marked on the number line (approximately half-way). Otherwise, this set was identical to the decimal stimuli-exponential anchor set.
- *Exponential Stimuli - Decimal Anchors*. These lines presented anchors in decimal notation and stimuli in exponential notation. Anchors were at 0 and 31,623.
- *Exponential Stimuli – Decimal Anchors Calibrated*. Before beginning trials within this set, participants were shown a single slide with the proper location of . $.500\text{x}10^{4.5}$ marked on the number line (approximately half-way). Otherwise, this set was identical to the exponential stimuli-decimal anchor set.

Procedure

Each participant completed several different sets of number line estimation tasks. Participants were randomly assigned to either of two conditions that varied the representational format of the endpoint anchors and to-be–estimated numbers (see Table 1). Condition 1 had 35 participants and condition 2 had 32 participants. All tasks were completed consecutively in one hour-long session.

Analyses, Results and Discussion

Analyses

The primary analyses involved comparisons of the fit of linear and logarithmic models to the median estimates for the numerical values. Specifically, we followed the method of Siegler and Booth (2006). First, we calculated the median estimate for each stimulus as generated by participants. Then the differences between median estimates and the number predicted by the best-fitting logarithmic and linear functions were compared via paired samples t-tests. The results of model comparisons for each set are summarized in Table 2.

Results and Discussion

Controls Performance on all control tasks were well accounted for by linear functions. These included the *Decimal 1-100, Decimal 1-1000, Decimal 0-32k, Decimals 16-32k, and All Exponential* sets. Because these sets were best fit by linear functions across conditions, the data were collapsed across conditions within each set (see Table 2). The fact that each of these sets yielded linear results supports the conclusion that participants' baseline performance was linear across the number ranges tested in the experiment. Importantly, the data showed that performance was linear for both decimal notation and exponential notation formats in the case for which the same format was used for both anchors and stimuli.

Decimal Stimuli - Exponential Anchors In contrast to control trial estimates, the estimates for this set were fit equally well by linear and logarithmic functions. In fact, though statistically insignificant, the absolute value of the variance explained was greater for the logarithmic model than for the linear model (see Table 2). It is doubtful, however, that this pattern of performance was due to the use of a logarithmic ruler, as performance in the control conditions demonstrated that participants were very capable of performing linearly in the same range in both standard and exponential notation. Instead, the compression seems to have been an artifact of confusion on the task (see Figure 1).

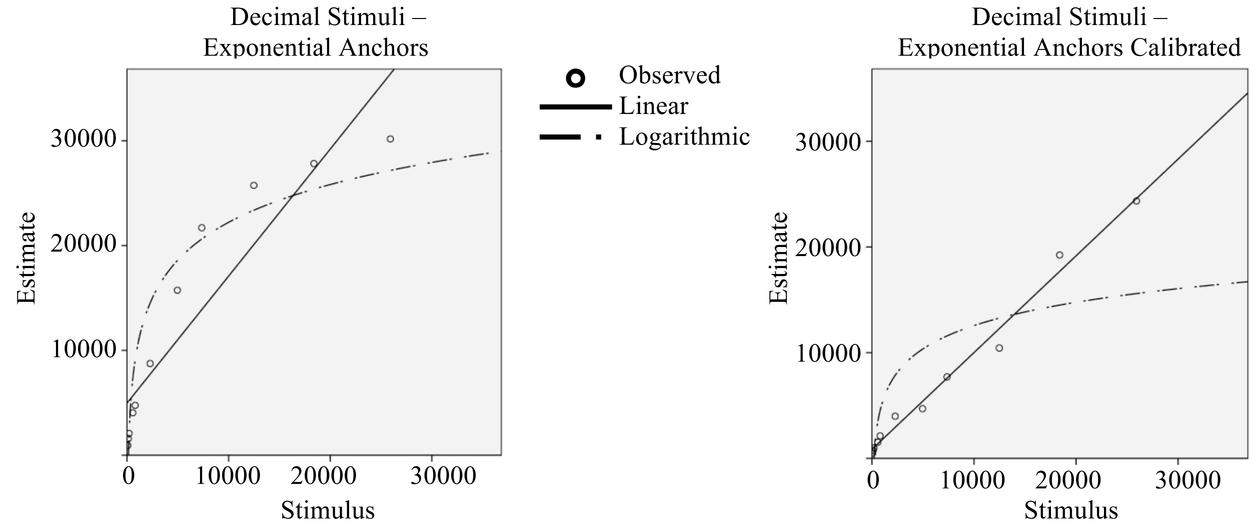

Figure 1: Performance before and after receiving information about a single data point

As predicted, there was a stark contrast between the linearity of performance for stimuli beneath 10,000 and the linearity of performance for stimuli above 10,000. For stimuli of value less than 10,000, aggregate performance was very linear with a slope well above 1, (R^2 = .98 vs. .83, linear slope = 2.72). This is confirmed by analysis of the individual data. Out of 32 participants, 22 (69%) were best fit by a linear function with a slope of 2 or greater in this range. Such high slopes are consistent with what would be predicted if participants thought the rightmost endpoint was approximately 10,000.

On the other hand, performance on stimuli greater than 10,000 was equally well fit by both linear and logarithmic functions.(R^2 = .99 vs. .99, linear slope = .33). Both types of functions fit well because there were only three data points for median estimates in this range (12,480, 18,390, and 25,920). Upon analyzing individual data, which included two data points for each estimate, it became clear that the data were not well behaved: the performances for 24 of 32 participants (78%) failed to fit a linear function with a slope different from zero in this range. This is particularly interesting given that performance on the decimal 16-32k set yielded linear performance in the same range. This suggests that the ostensibly logarithmic performance was due to confusion about the value of the uppermost anchor as opposed to the use of a logarithmic ruler. What looks like a compressive function was an artifact of how knowledge constraints affected adults' default procedure for executing the task.

Decimal Stimuli – Exponential Anchors Calibrated Performance with this group was far better fit with a linear function than with a compressive function (see Table 2). With the addition of a single slide, performance on this set changed dramatically (see Figure 1). The slide allowed participants to calibrate their estimates, making the value of the uppermost anchor meaningful. This is consistent with the hypothesis that the compressive performance was not indicative of the underlying representation, but was instead an artifact of confusion about the value of the uppermost anchor.

Exponential Stimuli - Decimal Anchors This set yielded linear performance, despite the compressive performance seen for the set using the converse notation (see Table 2). The slope of performance with this set is of particular interest. Even though the relative distances between the placements of the stimuli were linearly consistent, participants typically placed the large majority of the stimuli on the leftward third of the line. Indeed, the slope of .34 is what would be expected if participants thought that $.999 \times 10^{4.5}$ was equivalent to 10,000. When we rescaled the data in a way corresponding to the assumption that $.999 \times 10^{4.5}$ was actually equivalent to 10,000, the data yielded a slope of 1.10, again suggesting that performance was an artifact of mistaken assumptions about the uppermost anchor.

Exponential Stimuli - Decimal Anchors Calibrated With the addition of a single slide, performance on this set changed dramatically. Performance was still linear but the slope was greatly increased, yielding a value closer to 1, as opposed to the slope of .35 seen for the uncalibrated set (see Table 2). Note that it was not the form of the function that changed – it was linear for both sets – but its slope. Performance on this set corroborates our conclusions from the *Decimal Stimuli - Exponential Anchor* set: Divergence from linear performance with a slope of 1 is more the consequence of knowledge constraints than the result of different underlying representations of the numerical stimuli.

General Discussion

Adult performance in this study matched our predictions. Participants performed linearly on all tasks for which the values of the anchor points were unambiguous. However, when the (misleading) exponential notation was used, performance appeared more compressive.. Performance in this case was modeled at least as well by a logarithmic function as by a linear function. Because adults were linear over the same range in multiple notations, it seems that the compressive performance was an artifact of knowledge constraints. These performance patterns match those that should result from attempting to complete a proportion judgment task under conditions where necessary information (i.e., the values of anchors and to-be-placed stimuli) is incorrect or lacking.

These data are problematic for the stance that logarithmic performance on number line tasks is evidence that mental representations of numerical magnitude are logarithmically compressed. In particular, these results raise questions about whether children's logarithmic performance on number line placement tasks is due to them *only* having logarithmic representations available to them for a given number range. The current situation with adults is one case in which logarithmic *looking* performance was due to task constraints instead of reflecting a shift in the underlying representational system. Such considerations may similarly apply to children as well as to adults.

In sum, we offer these results as an existence proof that differences in performance on number line estimation tasks may not necessarily be indicative of fundamental differences in the format of people's underlying representations of numerical magnitude. Rather, they may be diagnostic of differences in people's understanding of what magnitudes are represented by a given numerical stimulus. This would explain Siegler and colleagues' consistent findings that linearity of performance on number line estimation tasks correlates with success in other areas of numerical ability (Ramani & Siegler, 2008; Siegler & Booth, 2004; Whyte & Bull, 2008). These tasks may not have been tracking changes in children's underlying numerical magnitude representations; they may instead have been gauging the extent to which children understood what values symbolic numbers represented. We suggest that such number line estimation tasks may prove useful in evaluating children's current level of understanding of the meaning of numbers, and agree with the thesis that they might serve as learning tools for helping children map decimal numbers and number words to appropriate mental magnitudes (Ramani & Siegler, 2008; Siegler, 2009). A series of studies is currently in progress to further investigate these possibilities. Particularly, we plan to investigate whether individual differences in children's number knowledge predict the points at which their number line placement estimates begin to appear logarithmic.

Acknowledgments

This research was made possible in part by support from the Moreau Academic Diversity Postdoctoral Fellowship Program of the University of Notre Dame. We would like to thank the members of Nicole McNeil's Cognition Learning and Development Lab for their help and support.

References

Barth, H. C., & Paladino, A. M. (2010). The development of numerical estimation: evidence against a representational shift. *Developmental Science, 14*, 125-135.

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

Clarke, B., & Shinn, M. R. (2004). A Preliminary Investigation into the Identification and Development of Early Mathematics Curriculum-Based Measurement. *School Psychology Review, 33*(2), 234-249.

Dehaene, S. (1997). *The number sense*. Oxford University Press New York:

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

Hollands, J. G., & Dyre, B. P. (2000). Bias in Proportion Judgments: The Cyclical Power Model. *Psychological Review*, *107*, 500-524.

Ramani, G. B., & Siegler, R. S. (2008). Promoting Broad and Stable Improvements in Low-Income Children's Numerical Knowledge Through Playing Number Board Games. *Child Development*, *79*, 375-394.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, *75*, 428-444.

Siegler, R. S., & Booth, J. L. (2005). Development of numerical estimation: A review. *Handbook of mathematical cognition*, 197–212.

Siegler, R. S., & Opfer, J. E. (2003). The Development of Numerical Estimation. *Psychological Science*, *14*, 237.

Siegler, R. S. (2009). Improving the Numerical Understanding of Children From Low-Income Families. *Child Development Perspectives*, *3*, 118-124.

Spence, I. (1990). Visual psychophysics of simple graphical elements. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 683-692.

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, *64*, 153-181.

Wright, B. (1991). What number knowledge is possessed by children beginning the kindergarten year of school. *Mathematics Education Research Journal, 3(1),* 1-16.