CHILD DEVELOPMENT

Child Development, November/December 2010, Volume 81, Number 6, Pages 1768–1786

How 15 Hundred Is Like 15 Cherries: Effect of Progressive Alignment on Representational Changes in Numerical Cognition

Clarissa A. Thompson
The University of Oklahoma

John E. Opfer
The Ohio State University

How does understanding the decimal system change with age and experience? Second, third, sixth graders, and adults (Experiment 1: N = 96, mean ages = 7.9, 9.23, 12.06, and 19.96 years, respectively) made number line estimates across 3 scales (0–1,000, 0–10,000, and 0–100,000). Generation of linear estimates increased with age but decreased with numerical scale. Therefore, the authors hypothesized highlighting commonalities between small and large scales (15:100::1500:10000) might prompt children to generalize their linear representations to ever-larger scales. Experiment 2 assigned second graders (N = 46, mean age = 7.78 years) to experimental groups differing in how commonalities of small and large numerical scales were highlighted. Only children experiencing progressive alignment of small and large scales successfully produced linear estimates on increasingly larger scales, suggesting analogies between numeric scales elicit broad generalization of linear representations.

The ratio structure of the decimal system—where ‘‘1’’ denotes a quantity 1/10 of 10, ‘‘10’’ a quantity 1/10 of 100, ‘‘100’’ a quantity 1/10 of 1,000, and so forth—may apply to an infinity of numbers, but even over a lifetime, experience of symbolic numbers is finite. And more, experiences are also systematically biased, with numerals in the first order of magnitude (1–9) appearing much more frequently than numerals in the second order (10–99), which occurs more frequently than in the third order (100–999), and so forth (Dehaene & Mehler, 1992). This cross-linguistic regularity in frequency of symbolic numbers—observed in languages as diverse as American English, Catalan, Dutch, French, Japanese, Kannada, and Spanish—has been observed across forms of notation (Arabic or written number words), for cardinal numbers (1, 2, 3) as well as ordinals (first, second, third), and in both text and speech.

The authors would like to thank the administration, teachers, students, and parents from the Worthington School District in Columbus, OH, as well as the Bentworth School District in Bentleyville, PA. Further, the authors would like to thank Julia Kennedy for her help in data collection, Chris Young for his contributions to the ‘‘change point’’ analyses in the General Discussion section, and anonymous referees who provided helpful comments on earlier versions of the manuscript. Portions of these data were presented at the 2007 meeting of the Cognitive Development Society as well as the 30th annual conference of the Cognitive Science Society.

Correspondence concerning this article should be addressed to Clarissa A. Thompson, Department of Psychology, The University of Oklahoma, 455 W. Lindsey Street, Dale Hall Tower, Room 727, Norman, OK 73019. Electronic mail may be sent to cat3@ou.edu.

This cross-linguistic regularity in frequency of symbolic numbers has two interesting implications for how children might develop their representations of numerical magnitude. The first developmental implication is that relative frequencies of numerals in the environment place a constraint on how broadly children extend their understanding of the decimal system. Specifically, children learning to map numerals to magnitudes would be expected to fail to realize that ratios that hold at smaller, more familiar orders of magnitude (e.g., relative magnitudes of 15 cherries and 100 cherries) extend to larger orders of magnitude (e.g., relative magnitudes of 1,500 and 10,000) with which they are less familiar.

The second implication is that if developing representations of large numerical magnitudes are constrained by poverty of input described by Dehaene and Mehler (1992), no inherent cognitive constraint would prevent rapid and broad representational changes from occurring in children. The crucial source of change in this account is children encountering information that relations among numbers at large orders of magnitude are similar to relations at small orders of magnitude (e.g., 150:1,000::15:100), much like relations at small orders of magnitude are similar to each other regardless of units involved (e.g., 15 cherries:100 cherries::15 pears:100 pears). If true, such findings

© 2010 The Authors
Child Development © 2010 Society for Research in Child Development, Inc.
All rights reserved. 0009-3920/2010/8106-0012

would be scientifically important because it could reconcile two sets of seemingly contradictory findings—slow rate of representational changes observed in cross-sectional studies (e.g., Siegler & Opfer, 2003; logarithmic-to-linear switch in numeric representations between second and fourth grade on 0–1,000 number line problems) and one-trial representational changes observed in microgenetic studies (e.g., Opfer & Siegler, 2007; second graders' adoption of a linear representation after maximally discrepant feedback is presented). That is, representational change in the domain of numbers may be driven by children's ability to draw an analogy between small and large orders of magnitude, and previous research has indicated that direct feedback may highlight this analogy of the decimal system for children (Opfer & Siegler, 2007).

Previous evidence supporting the first conjecture about children's developing understanding of this property of the decimal system—as well as potential for improvement coming from comparison of small and large orders of magnitude—comes from interindividual variability in magnitudes children associate with a given number. When asked to place numbers on a number line flanked by 0 and 100, for example, second graders provided Siegler and Opfer (2003) with a set of estimates that were highly accurate and increased linearly with numeric value. When asked to place the same numbers on a number line flanked by 0 and 1,000, however, these same children provided estimates that were much less accurate, with estimates increasing logarithmically with numeric value. Thus, across contexts, the magnitude ''15'' was estimated as having different relations to 100 and 1,000. Moreover, this variability appeared to reflect variability in children's thinking about numeric magnitude rather than ''noise'' attributable to measurement error: Given no feedback, children's numerical magnitude estimates on a number line are highly stable from trial to trial (Opfer & Siegler, 2007; Opfer & Thompson, 2008; Thompson & Opfer, 2008).

To address these two implications more directly, Experiment 1 was designed to test how broadly children spontaneously apply linear representations on number lines with large numeric anchors (e.g., 0–1,000, 0–10,000, and 0–100,000). To test whether analogy could serve as a mechanism of developmental change across numeric representations, Experiment 2 examined whether progressive alignment—a means of fostering analogies in young children (Gentner, Loewenstein, & Hung, 2007; Kotovsky & Gentner, 1996)—led numeric representations used at small numerical scales (0–100) to be generalized to progressively larger numerical scales (0–1,000, 0–10,000, and 0–100,000) simply by highlighting perceptual similarities across these numeric contexts.

In the next sections, we will detail (a) development of numeric representations across the life span, (b) analogy as a mechanism of developmental change across numeric representations, and (c) more specific empirical questions that were examined in the present studies.

Development of Numerical Representations

Children normally improve their understanding about magnitudes denoted by symbolic numerals across a wide range of tasks (e.g., number line estimation, number categorization, magnitude judgments, etc.). Development of numerical representations appears to occur iteratively, with parallel developmental changes occurring over many years and across many contexts (Siegler, Thompson, & Opfer, 2009). For example, sixth graders' estimates on 0–1,000 number lines increase linearly (Siegler & Opfer, 2003), whereas second graders' estimates increase logarithmically (Opfer & Siegler, 2007; Opfer & Thompson, 2008; Siegler & Opfer, 2003; Thompson & Opfer, 2008). On 0–100 number lines, second graders' estimates increase linearly (Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007; Geary, Hoard, Nugent, & Byrd-Craven, 2008; Siegler & Booth, 2004; Siegler & Opfer, 2003), whereas kindergartners' estimates increase logarithmically (Siegler & Booth, 2004). Asked to estimate 0–10 chips from a pile of chips, kindergartners' estimates increase linearly, but preschoolers' estimates increase logarithmically (Opfer, Thompson, & Furlong, 2010). Further evidence for logarithmic-to-linear shifts in numerical magnitude representations have also been discovered when (a) children were asked to estimate candies in a container, (b) estimate salaries given in fractional notation, (c) provide answers to arithmetic problems, (d) make measurements of novel units, and (e) categorize and compare symbolic numbers (Booth & Siegler, 2006; Laski & Siegler, 2007; Opfer & DeVries, 2008; Opfer & Thompson, 2008; Thompson & Opfer, 2008).

Children's initial expectations that numerical magnitudes increase logarithmically are theoretically interesting because these expectations are consistent with Fechner's law. That is, just as sensations increase logarithmically with stimulus intensity (Fechner's law), representations of numeric magnitude also increase logarithmically with actual value (Dehaene, 2007). For instance, the difference between 1 and 10 cherries seems larger (or is more quickly detected) than the difference between 101 and 110 cherries, much as these numbers would be spaced on a logarithmic ruler (Dehaene, Dehaene-Lambert, & Cohen, 1998). This logarithmic representation of numerical magnitude is not unique to young children making estimates on number lines but is widespread across tasks (estimation of set sizes, numeric comparisons, random number generation, and economic games), across species (pigeons, rats, nonhuman and human primates), and across age groups (infants, children, and time-pressured adults; Banks & Hill, 1974; Brannon, 2005; Feigenson, Dehaene, & Spelke, 2004; Furlong & Opfer, 2009; Gallistel & Gelman, 1992; Moyer & Landauer, 1967; Roberts, 2005; Xu & Spelke, 2000).

Children's learned expectations that numerical magnitudes increase linearly are not just theoretically interesting but also have practical effects on their school achievement. For instance, linearity of children's number line estimates correlates strongly with how quickly they compare magnitudes (e.g., is 5 or 7 greater; Laski & Siegler, 2007), their ability to learn solutions to unfamiliar addition problems (Booth & Siegler, 2008), and their overall scores on mathematics achievement tests (Booth & Siegler, 2006; Siegler & Booth, 2004). Playing numeric board games (akin to Chutes and Ladders) has also been shown to impact preschoolers' ability to compare numerical value and to produce accurate answers to arithmetic problems (Griffin, Case, & Siegler, 1994; Ramani & Siegler, 2008; Siegler & Ramani, 2008, 2009). Thus, not only does the development of linear representations of numerical magnitude present an interesting scientific problem, it is also one with important educational consequences, an importance underlined by recommendation of the National Mathematics Advisory Panel to embrace practices that foster ''number sense,'' including ability ''to estimate orders of magnitude'' (U.S. Department of Education, 2008, p. 18).

Analogy as a Mechanism of Developmental Change Across Numeric Representations

By what mechanism might children abandon use of a logarithmic representation of numerical magnitude for use of a linear representation? Across many areas of cognitive development, an important mechanism for representational change is analogy, where elements in a target domain may share structural relations with a base domain despite lack of perceptual overlap among elements in the target and base domains (Chen & Klahr, 1999; Gentner et al., 1997; Holyoak & Thagard, 1995; Opfer & Siegler, 2004, 2007).

Although young children may initially have difficulty solving analogical reasoning problems (see Abdellatif, Cummings, & Maddux, 2008, for a recent review of the factors affecting the development of analogical reasoning in young children), a number of factors can improve analogical reasoning in children such as: (a) asking children to solve problems with familiar examples or familiar relations (e.g., loaf of bread:single slice of bread::lemon:___; Goswami & Brown, 1990), (b) providing direct instruction on higher order relations through the use of relational language (e.g., ''Daddy, Mommy, and Baby'' for sizes big, medium, and small; Rattermann & Gentner, 1998), (c) explaining classic analogies of the format A:B::C:D (e.g., noting the relation between A and B and the similarity between A and C and B and D; Alexander, Willson, White, & Fuqua, 1987), (d) training children to solve a base problem and then asking them to transfer this knowledge to solve a related target problem (e.g., children are told a story about a genie who moves jewels from one bottle to another by rolling a ''magic carpet'' into a tube shape, and then the children are asked to solve the problem of moving gumballs from one bowl to another; Holyoak, Junn, & Billman, 1984), and (e) comparing across various instances (e.g., Gentner & Namy, 1999).

Structure-mapping theory (Gentner, 1983) guides our understanding of children's analogical reasoning ability in the current experiments. According to structure-mapping theory, children form analogies by aligning representational elements between a base and target domain. This alignment process facilitates transfer of information from base to target through children's comparison of surface-level features. This comparison process leads to subsequent highlighting of common underlying relational structure shared by base and target (Kurtz, Miao, & Gentner, 2001).

In progressive alignment (Gentner et al., 2007; Kotovsky & Gentner, 1996)—a specific instantiation of structure-mapping theory principles—comparisons made between highly similar elements can promote subsequent analogical matches of lower overall similarity. Analogical matches are thus promoted because the process of aligning surface-level elements highlights common underlying relational structure. The alignment process makes this structure more salient and easily mapped from base to target domain even when participants are presented with less surface-similar matches.

According to Kotovsky and Gentner (1996), progressive alignment acts as a mechanism of representational change by allowing children to make similarity comparisons over concrete, perceptual similarities (e.g., monotonic increase in size across differently shaped stimuli). Then, these similarity comparisons facilitate children's ability to notice higher order relational commonalities across stimuli possessing fewer surface-level features in common (e.g., increase in size as compared to saturation of color across differently shaped stimuli). Thus, progressive alignment allows children to recognize ''richer and deeper'' abstract relational similarity uniting their mental representations that may not have been immediately apparent before similarity comparisons were made (Kotovsky & Gentner, 1996).

Our general perspective on representational change, drawn from computational models of cognition (Doumas, Hummel, & Sandhofer, 2008; Gentner, 1983; Hummel & Holyoak, 2003), historical changes in scientific concepts (Gentner et al., 1997; Holyoak & Thagard, 1995), and microgenetic studies of children's concept learning (Opfer & Siegler, 2004, 2007) immediately suggested analogy as a candidate mechanism for development of numerical representations. This hypothesis seemed especially likely to be true for number line estimates, where structure of elements in a number line problem (e.g., relation between right anchor and number to be estimated) offers a systematic regularity over orders of magnitude (e.g., 150:1,000::15:100).

Analogy was also suggested as a candidate mechanism of representational change for second-grade students in a recent number line training study conducted by Opfer and Siegler (2007). Although these second-grade students produced a logarithmic series of estimates on 0–1,000 number lines during pretest, when children received feedback on the placement of 150 (maximally discrepant point between logarithmic and linear functions forced to pass through 0 and 1,000), children immediately generated a linear series of estimates for all other numbers in the 0–1,000 numeric scale that were tested. This abrupt and broad generalization of learning—characteristic of analogical mapping more broadly (Chen & Klahr, 1999; Gentner, Holyoak, & Kokinov, 2001; Holyoak & Thagard, 1995; Opfer & Siegler, 2004)—led Opfer and Siegler to speculate that children had improved their performance by mapping commonalities between the less familiar, larger scale (0–1,000) to the more familiar, smaller scale (0–100) where children already possessed a linear representation.

In this article, we provided a novel test of Opfer and Siegler's (2007) hypothesis that analogy provides a mechanism for developmental changes in representations of numerical magnitude. Our test was unique in two important respects. First, if children can draw an analogy between 0–100 and 0–1,000 number line problems, feedback that Opfer and Siegler provided in the 0–1,000 scale would not be necessary for children to improve their performance on those problems. To test this issue, our strategy for eliciting analogies from children was to manipulate factors that would lead to comparisons of number line problems, and we otherwise avoided giving children feedback on target problems. Second, if children can extract general structural relations hypothesized by Opfer and Siegler (i.e., 150:1,000::15:100), they should generalize far outside the training space (e.g., from a 0–100 number line to a 0–100,000 number line), an issue that Opfer and Siegler did not address.

Further, we have proposed that possession of a numeric analogy between familiar, small numeric scales (0–100) and unfamiliar, larger numeric scales (0–1,000 and larger) can prompt children to scale up a linear representation of number to unfamiliar scales. Thus, familiarity with numeric context (0–100) plays an important role in our explanation of analogy as a mechanism of representational change, but another model of children's estimation performance, the segmented linear model (Ebersbach, Luwel, Frick, Onghena, & Verschaffel, 2008), suggests that a very different type of familiarity affects children's estimates.

According to Ebersbach et al.'s (2008) segmented linear model, children's familiarity with numbers (indexed by their counting ability within a particular numeric scale) is a good predictor of shape of the participants' mental number line (e.g., linear or logarithmic). That is, if participants are familiar with the numeric context in which they are asked to make estimates, the best fitting function of participants' number line estimates should be linear with a greater slope than the scale with which they are unfamiliar, thereby leading to a pattern of estimates that appears logarithmic. Within this account, apparent fit of the logarithmic model is illusory, whereas a segmented linear model, consisting of two simple linear regression functions that intersect at a ''change point'' provides a better model fit. Most impressively, ''change point'' in this segmented linear model was correlated with children's counting range, suggesting the apparent fit of the logarithmic function may be a by-product of children's familiarity. In this article, we investigate whether the assumptions of the segmented linear model hold in larger numeric scales (e.g., 0–10,000 and 0–100,000).

Issues Examined in Present Studies

The central purpose of the present studies was to examine how broadly children spontaneously apply linear representations on number lines with large numeric anchors (e.g., 1,000, 10,000, and 100,000; Experiment 1) and to determine whether aligning structure of small and large numeric scales will prompt children to generalize linear representations more broadly than they would without such alignment (Experiment 2).

The purpose of Experiment 1 was to determine age differences in underlying representations that children and adults used when making estimates on number lines of increasing numerical magnitude and track the developmental trajectory of participants' linearization of estimates on number lines of increasing magnitudes. To investigate these age differences in numerical representations, we examined second-, third-, and sixth-grade students' number line estimation performance in large numerical contexts (e.g., 0–1,000, 0–10,000, and 0–100,000) prior to children receiving any corrective feedback from the experimenter. These children's results were then compared to the performance of college-aged adults who were asked to complete the same task. From the results of Experiment 1, we also hoped to identify an age group where there was a large potential for generalizing a linear representation of numbers to much larger orders of magnitude. Experiment 1 also allowed us to investigate whether the assumptions of Ebersbach et al.'s (2008) segmented linear model held at larger orders of magnitude (e.g., 0–10,000 and 0–100,000 scales).

In Experiment 2, we attempted to (a) determine how widely children generalized linear representations of numerical magnitudes after they were given feedback on the correctness of their estimates on a smaller scale and (b) determine whether alignment of small and large numeric scales was sufficient to produce generalization to larger numerical contexts. To meet our goals for Experiment 2, we brought number lines for large scales (0–1,000, 0–10,000, and 0–100,000) into progressive alignment with scales that children already represented linearly (0–100). The alignment procedure we used is considered ''progressive'' in two senses. First, all children's comparisons were supported by increasing perceptual similarity of small (training problems) and large scales (generalization problems) by matching color of units and orders of magnitude. Second, in our progressive alignment and multiple exemplars conditions, children were able to compare 0–100 problems with 0–1,000, 0–10,000, and 0–100,000 problems; in contrast, children in the no alignment condition were not given this opportunity. Finally, to test for representational change, we examined numerical estimates on a 0–1,000 posttest, where children were not given feedback, perceptual support, or the opportunity to compare problems to smaller scales.

Experiment 1: Long-Term Changes in Representations of Large Numerical Magnitudes

In Experiment 1, we investigated long-term changes in children's estimates of large numerical magnitudes (e.g., 940, 9,400, and 94,000) by examining estimates of second graders, third graders, sixth graders, and adults on 0–1,000, 0–10,000, and 0–100,000 number lines. Previously, number line estimation results have shown parallel developmental changes at different ages. Thus, a logarithmic-to-linear shift occurs between kindergarten and second grade on 0–100 number lines (Siegler & Booth, 2004), and between second and fourth grade for estimates on 0–1,000 number lines (Siegler & Opfer, 2003). In this study, we examined whether yet another logarithmic-to-linear shift occurs among children older than second graders on 0–10,000 and 0–100,000 number lines. This issue was important because findings of parallel developmental changes at older ages would suggest that for many years children face the problem of not realizing when to scale up their linear representation and would thereby raise the scientific issue of how they might make that realization.

Method

Participants. Participants were 24 second graders (mean age = 7.9, *SD* = 0.33; 14 girls and 10 boys), 24 third graders (mean age = 9.23, *SD* = 0.36; 11 girls and 13 boys), 24 sixth graders (mean age = 12.06, *SD* = 0.36; 12 girls and 12 boys), and 24 college-aged adults (mean age = 19.96, *SD* = 1.8; 16 women and 8 men). Children were recruited from elementary schools in largely European American, middle-class suburbs surrounding two large metropolitan cities in the United States. Adults were recruited from an introductory psychology course at a large university in the same city. Children participated in return for a small prize (e.g., a sticker), whereas adults received credit toward their introductory psychology course. One female graduate student and one female research assistant served as experimenters.

Design and procedure. Children and adults were given a number line estimation task that consisted of a line flanked by two hatch marks where the left hatch mark was labeled ''0,'' and right hatch mark was labeled with either ''1,000,'' ''10,000,'' or ''100,000'' depending on experimental condition (see Figure 1). Magnitude of overall scale (0–1,000, 0–10,000, and 0–100,000) was a between-subjects variable to guard against order effects. Across all conditions, participants were asked to estimate the position of a third number that appeared above the midpoint of the number line by making a hatch mark through the line. To-be-estimated numbers were chosen to maximize discriminability of logarithmic and linear functions (i.e., by oversampling at the low end of the scale) and to minimize influence of specific knowledge (e.g., 500 is halfway between 0 and 1,000).

Participants were presented with 1 number line estimation problem per page, which ensured that participants were unable to reference their previous estimates. Participants from each age group were randomly assigned to one of three conditions where they completed 10 number line estimation problems, presented in random order and without feedback from the experimenter: (a) in the 0–1,000 condition, participants estimated the magnitudes 20, 50, 80, 110, 150, 250, 490, 610, 730, and 940; (b) in the 0–10,000 condition, participants estimated the magnitudes 200, 500, 800, 1,100, 1,500, 2,500, 4,900, 6,100, 7,300, and 9,400; and (c) in the 0–100,000 condition participants estimated the magnitudes 2,000, 5,000, 8,000, 11,000, 15,000, 25,000, 49,000, 61,000, 73,000, and 94,000.

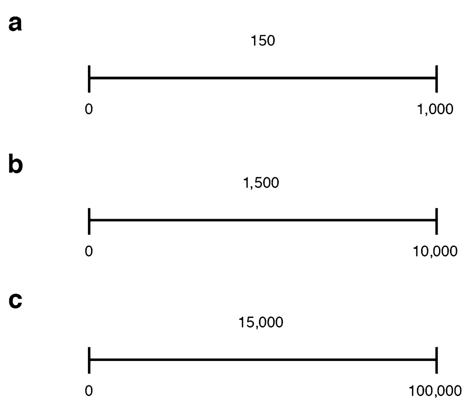

Figure 1. Experiment 1: Illustration of stimuli: (a) 0–1,000 scale, (b) 0–10,000 scale, and (c) 0–100,000 scale.

Results and Discussion

We first examined accuracy of numerical estimates across three experimental conditions. To measure accuracy, we first converted the participant's hatch mark to a numeric value by taking the proportion of the line indicated by the participant (i.e., linear distance from ''0'' mark to participant's hatch mark, divided by total length of line) and then multiplied this proportion by the right numerical anchor (1,000, 10,000, or 100,000). Then, percent absolute error of each participant's error (0%–100%) was calculated by taking the mean absolute difference between each of the participants' estimated values and the actual values divided by the total scale (1,000, 10,000, or 100,000). Finally, accuracy scores were computed by subtracting percent absolute error from 100%.

To analyze long-term changes in estimation accuracy, we conducted a 4 (age group: second graders, third graders, sixth graders, adults) × 3 (scale: 1,000, 10,000, 100,000) analysis of variance (ANOVA) on accuracy scores. As expected, accuracy increased with age, $F(3, 95) = 27.43$, $p < .0001$, $\eta^2 = .74$, and decreased with scale, $F(2, 95) = 7.51$, $p < .001$, $\eta^2 = .13$. These two main effects were also qualified by a significant Age Group × Scale interaction, $F(6, 95) = 2.38$, $p < .05$, $\eta^2 = .13$ (see Table 1). Although third and sixth graders outperformed second graders on 0–1,000 and 0–10,000 number lines, third and sixth graders failed to outperform second graders on 0–100,000 number lines. Thus, age group differences in estimation accuracy depended on numerical context (cf. Siegler & Opfer, 2003), with larger age differences present on 0–1,000 and 0–10,000 number lines than were present on 0–100,000 number lines.

To examine whether age differences in estimation accuracy for 0–10,000 and 0–100,000 number lines were linked to the logarithmic-to-linear shift seen in children's estimates on 0–1,000 number lines (Opfer & Siegler, 2007; Opfer & Thompson, 2008; Siegler & Opfer, 2003; Thompson & Opfer, 2008) and 0–100 number lines (Booth & Siegler, 2006, 2008; Laski & Siegler, 2007; Siegler & Booth, 2004; Siegler & Mu, 2008), we regressed participants' median estimates against the to-be-estimated number, and we compared the fit of the best fitting linear and logarithmic regression functions across the three age groups and scales (see Figure 2). Consistent with previous usage (e.g., Siegler & Opfer, 2003; Thompson & Opfer,

Table 1
Accuracy of Numerical Estimates by Age Group and Condition

	0–1,000	0–10,000	0–100,000
Age	$F(3, 31) = 10.92$, $p < .0001$ Adults, M = 96.2% > Sixth, M = 91%, d = 1.36 Adults, M = 96.2% > Third M = 90.5%, d = 1.12 Adults, M = 96.2% > Second, M = 80.3%, d = 2.94 Sixth, M = 91% > Second, M = 80.3%, d = 1.72 Third, M = 90.5% > Second, M = 80.3%, d = 1.45	$F(3, 31) = 28.96$, $p < .0001$ Adults, M = 96.7% > Sixth, M = 89.2%, d = 1.42 Adults, M = 96.7% > Third, M = 85.6%, d = 1.93 Adults, M = 96.7% > Second, M = 67.9%, d = 5.91 Sixth, M = 89.2% > Second, M = 67.9%, d = 3.03 Third, M = 85.6% > Second, M = 67.9%, d = 2.39	$F(3, 31) = 6.36$, $p < .01$ Adults, M = 96.3% > Sixth, M = 80.4%, d = 1.84 Adults, M = 96.3% > Third, M = 77.7%, d = 2.65 Adults, M = 96.3% > Second, M = 73.8%, d = 1.96

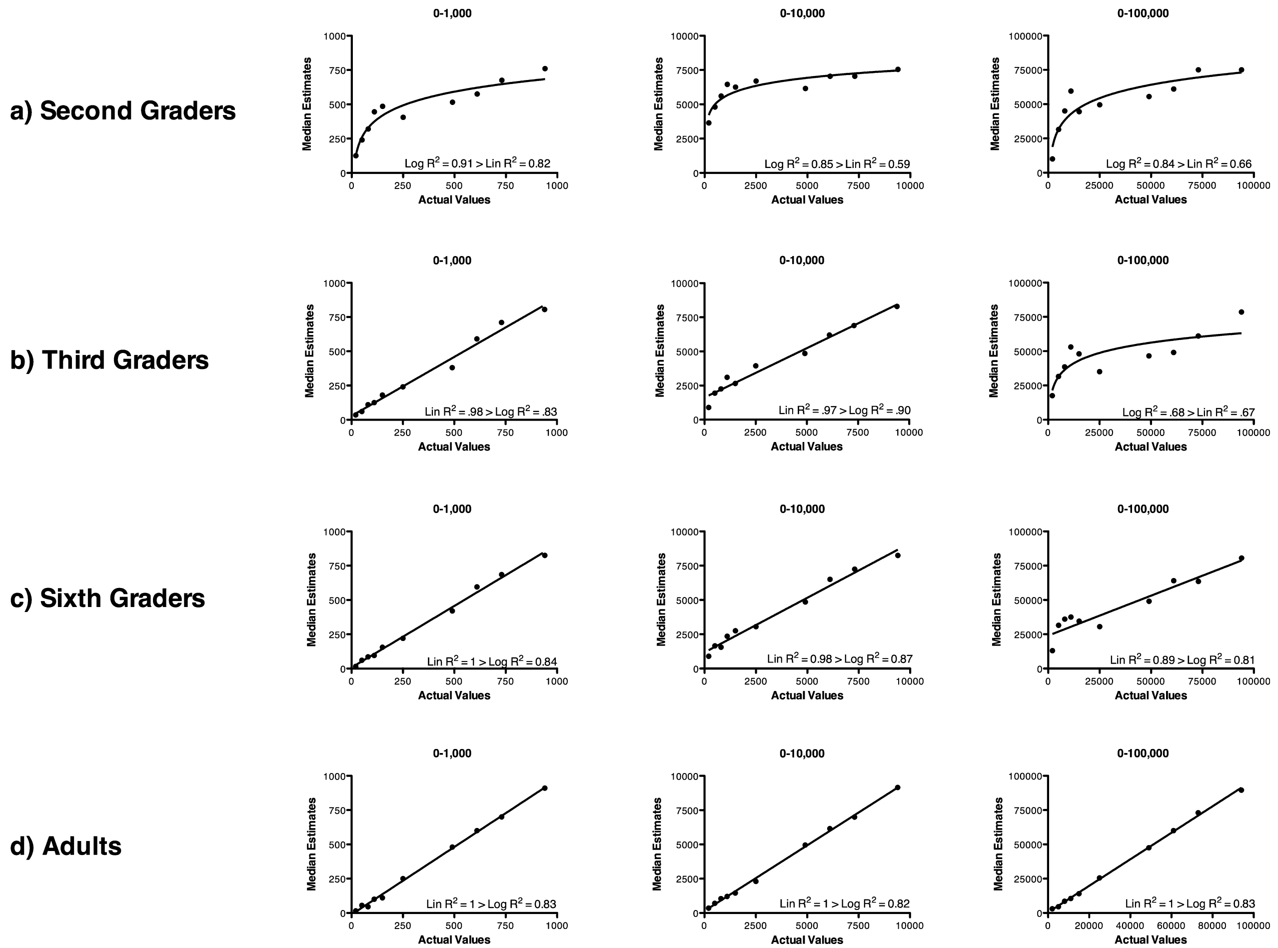

Figure 2. Experiment 1: Median estimates of second graders, third graders, sixth graders, and adults on 0–1,000, 0–10,000, and 0–100,000 number lines.

2008), median estimates were used to minimize the impact of potential outliers at the group level (although similar logarithmic and linear fits were obtained when we used mean estimates in Experiments 1 and 2). Across all three scales, second graders' median estimates were better fit by logarithmic than linear regression functions (0–1,000: logarithmic R^2 = .91 > linear R^2 = .82; 0–10,000:

logarithmic R^2 = .85 > linear R^2 = .59; 0–100,000: logarithmic R^2 = .84 > linear R^2 = .66). Third graders' median estimates, however, were better fit by linear than logarithmic functions on 0–1,000 and 0–10,000 scales (0–1,000: linear R^2 = .98 > logarithmic R^2 = .83; 0–10,000: linear R^2 = .97 > logarithmic R^2 = .90), but their median estimates on 0–100,000 number lines were better fit by logarithmic than by linear regression functions (0–100,000: logarithmic R^2 = .68 > linear R^2 = .67). Finally, median estimates on all three number lines were better fit by linear than by logarithmic regression functions for sixth graders (0–1,000: linear R^2 = 1 > logarithmic R^2 = .84; 0–10,000: linear R^2 = .98 > logarithmic R^2 = .87; 0–100,000: linear R^2 = .89 > logarithmic R^2 = .81) and adults (0–1,000: linear R^2 = 1 > logarithmic R^2 = .83; 0–10,000: linear R^2 = 1 > logarithmic R^2 = .82; 0–100,000: linear R^2 = 1 > logarithmic R^2 = .83).

To determine if the results at the group level held at the individual level, we next regressed each individual participant's estimates against the actual number using logarithmic and linear regression functions, and we used logistic regression to examine the odds of providing linear series of estimates as a function of age (see Figure 3). On 0–1,000 number lines, odds of generating linear estimates tended to increase with age, $\hat{\beta} = .72$, z = 1.90, Wald(1, N = 32) = 3.60, p = .058, indicating that with each year of age children were 2.05 times as likely to generate linear series of estimates. On 0–10,000 number lines, odds of generating linear estimates also increased with age, $\hat{\beta} = .68$, z = 2.28, Wald(1, N = 32) = 5.2, $p < .05$, indicating that with each year of age children were 1.97 times as likely to generate linear series of estimates. Finally, on 0–100,000 number lines, odds of generating linear estimates increased with age, $\hat{\beta} = .35$, z = 2.32, Wald(1, N = 32) = 5.4, $p < .05$, indicating that with each year of age children were 1.41 times as likely to generate linear series of estimates.

In summary, estimation accuracy increased with age, with age differences in accuracy being smaller on progressively larger numerical scales. The same pattern was also evident in the proportion of children producing linear series of estimates. Within each numerical scale, generation of linear series of estimates increased with age, but the overall proportion of children generating linear series of estimates again decreased with progressively larger scales. Thus, taken with previous results on numerical estimation in 0–10 and 0–100 contexts, the results of Experiment 1 point to a problem in numerical estimation that is faced to some extent by children from preschool to early adolescence: How to recognize that the linear representation used for small numbers is also appropriate to use for large numbers?

Experiment 2: Effect of Progressive Alignment on Representational Change

In Experiment 2, we investigated the effect of progressive alignment on children's application of linear representations to large numerical scales. Results of Experiment 1 suggested that second graders would serve as ideal participants: Second graders typically represent numbers as increasing linearly on 0–100 number lines and logarithmically on 0–1,000 number lines (Siegler & Booth, 2004; Siegler & Opfer, 2003), thereby making an analogy between the two scales potentially effective in improving representations for larger numbers. To test this idea, we

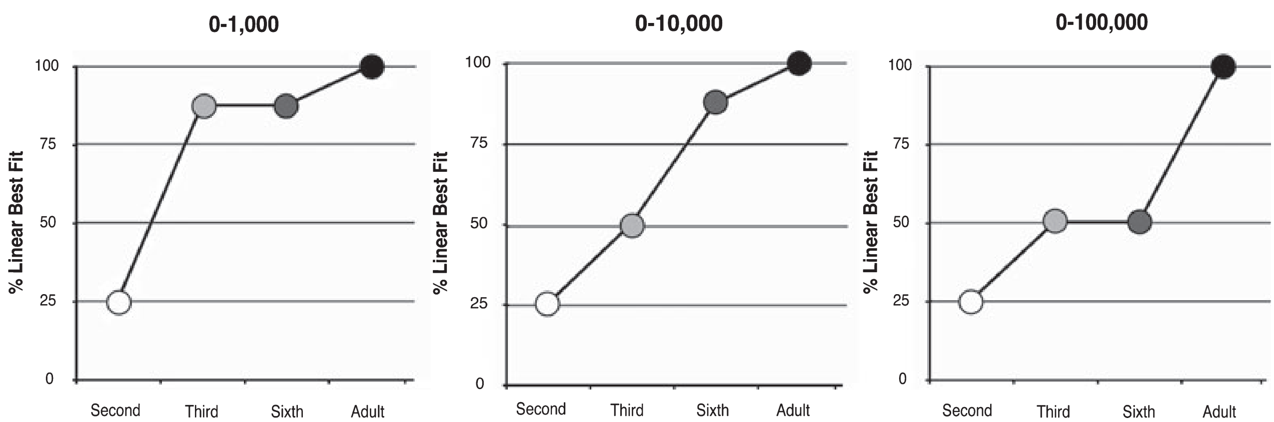

Figure 3. Experiment 1: Percentage of second graders, third graders, sixth graders, and adults who were best fit by the linear regression function on 0–1,000, 0–10,000, and 0–100,000 number lines.

examined effects of three training procedures to elicit linear estimates from second graders on 0–1,000, 0–10,000, and 0–100,000 number lines. In all three procedures, children were given feedback when making estimates on 0–100 number lines. The critical difference between procedures, however, was whether and how 0–100 number lines were aligned with 0–1,000, 0–10,000, and 0–100,000 number lines. Our central hypothesis was that progressive alignment of small and large number lines would lead children to change their estimates of numerical magnitude on large number lines, despite being given no direct feedback that they should do so.

We hypothesized that progressive alignment would promote representational change in the following way. Participants should first make surface-level comparisons across 0–100 training problems and note that a particular numerosity should be placed in the same location on the number line regardless of whether the numerosity is presented in the context of a blank 0–100 number line, a 0–100 pear line, a 0–100 cherry line, or a 0–100 carrot line. Then, we hypothesized that participants should note color of units (e.g., pears, cherries, carrots) corresponding to differing orders of magnitude (0–1,000 with one green zero, 0–10,000 with two red zeros, and 0–100,000 with three orange zeros, respectively), and this should subsequently highlight the underlying decimal system (15 cherries:100 cherries::1,500:10,000).

Method

Participants. Participants were 46 second graders (mean age = 7.78, *SD* = 0.41) recruited from elementary schools in largely European American, middle-class suburbs surrounding two large metropolitan cities in the United States. There were 20 girls and 26 boys that participated in return for a small prize (e.g., a sticker). One female graduate student served as experimenter.

Design and procedure. Children estimated the placement of numbers on number lines across three phases of the experiment: training, generalization, and posttest (see Figure 4). Number lines were flanked by two hatch marks, the left hatch mark was labeled ''0,'' and right hatch mark was labeled ''100,'' ''1,000,'' ''10,000,'' or ''100,000.'' On each trial, children estimated the position of a number (one per number line) by making a hatch mark through the line. To-be-estimated numbers (2, 5, 8, 11, 15, 25, 49, 61, 73, 94, and multiples thereof) were chosen to reduce influence of specific knowledge (e.g., that 50 is half of 100) and to oversample the low end of the scale to maximize discriminability of logarithmic and linear functions.

Figure 4 illustrates training, generalization, and posttest phases for three experimental conditions: no alignment, multiple exemplars, and progressive alignment. During training (Figure 4, left column), all participants received corrective feedback in the 0–100 scale that indicated how close to or far from the actual location of the to-be-estimated number their hatch marks were to ensure children possessed a linear representation in this scale (see Opfer & Siegler, 2007, for a more detailed description of the feedback procedure). We reasoned that if children possessed a linear representation in the 0–100 scale, participants across all experimental groups should be equally likely to bootstrap their linear representation to larger numerical scales (0–1,000, 0–10,000, and 0–100,000).

All training problems specified units (i.e., green pears, red cherries, and orange carrots) and were completed in the following order: blank number line, pear number line, cherry number line, then carrot number line. After children completed these training problems, they were told that their estimates did not differ much over different units (e.g., ''It does not matter if you see pears, cherries, carrots, or nothing at all after the number 15, if you see the number 15, you should make your mark right here'').

In the generalization phase (Figure 4, middle column), we highlighted similarity of generalization and training problems (that color of units matched color of zeros in right numerical anchor) and told participants to ''try some more problems just like the ones you just finished'' as the experimenter drew an arrow from the relevant training problem to its related generalization problem. During generalization, participants made estimates in the 0–100, 0–1,000, 0–10,000, and 0–100,000 scales without corrective feedback for the numerosity on which they were just trained.

Figure 4 illustrates training and generalization for the number 15. Participants completed one training problem at a time (e.g., 15 on a 0–100 number line, 15 pears on a 0–100 pear line, 15 cherries on a 0–100 cherry line, and 15 carrots on a 0–100 carrot line), and then completed generalization problems in the 0–100 through 0–100,000 scale (e.g., 15 on a 0–100 number line, 150 on a 0–1,000 number line, 1,500 on a 0–10,000 number line, and 15,000 on a 0–100,000 number line). This training and generalization procedure was repeated with the other nine numerosities we investigated for a total of 40 training trials and 40 generalization trials.

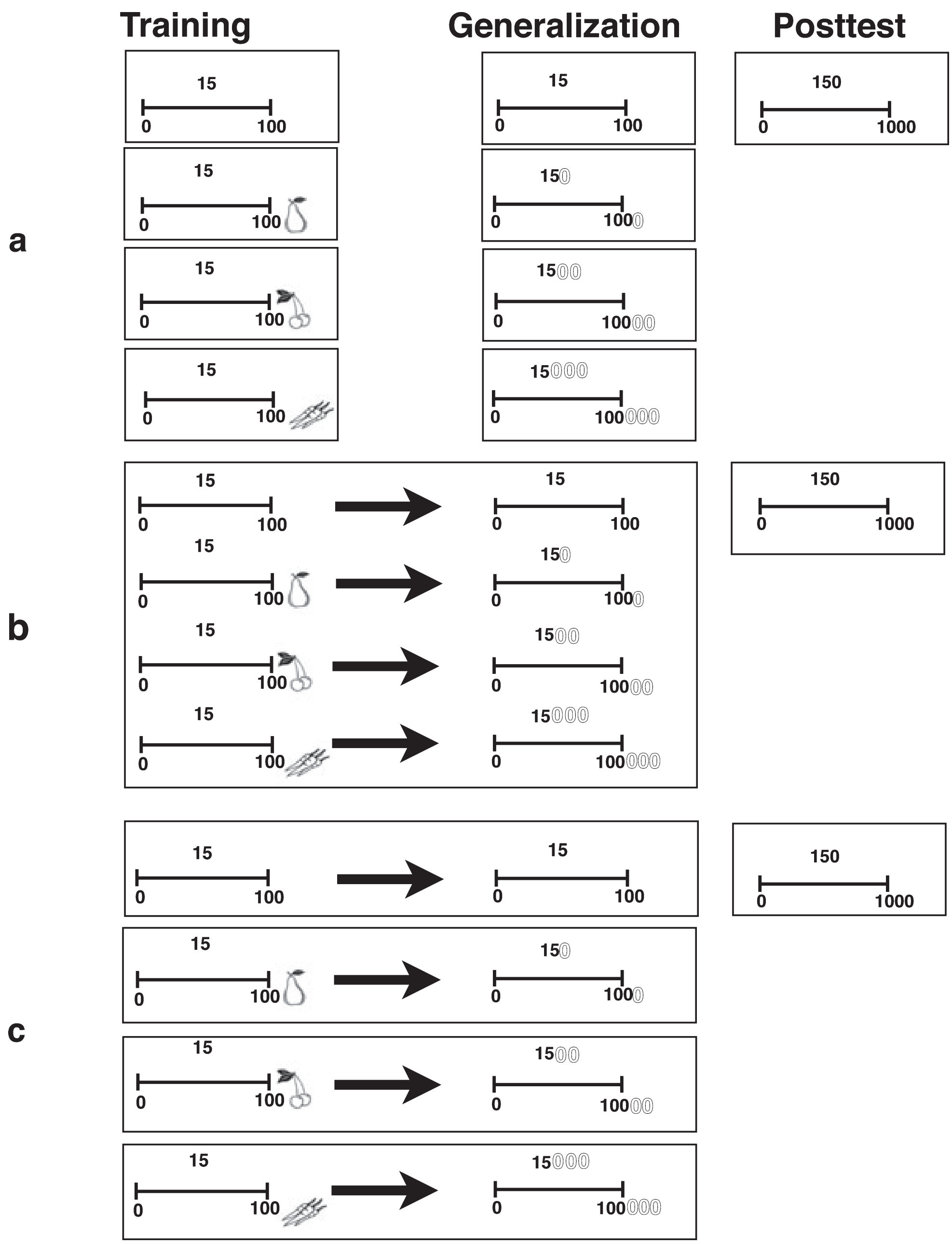

Figure 4. Experiment 2: Illustration of stimuli: (a) no alignment condition, (b) multiple exemplars condition, and (c) progressive alignment condition.

Note. Units in 0–100 training problems (i.e., pears, cherries, and carrots) were colored to correspond to the final 1–3 zeros in 0–1,000, 0–10,000, and 0–100,000 generalization problems (e.g., 3 carrots and the final three zeros in 100,000 were colored orange). Participants in the no alignment condition were presented one training or generalization problem per page, so these participants could not compare across training and generalization problems. Participants in the multiple exemplars condition were presented all training and generalization problems on one page, so these participants could compare training to other training problems, generalization to other generalization problems, or training to generalization problems. Participants in the progressive alignment condition were presented with one training–generalization pair at a time, so these participants could compare across relevant training–generalization pairs only.

Following completion of the training and generalization procedure detailed earlier, participants completed a 10-problem posttest phase (Figure 4, right column). Participants completed the same 10 problems administered for the 0–1,000 generalization phase (e.g., 20, 50, 80, 110, 150, 250, 490, 610, 730, 940). Like the generalization phase, participants did not receive corrective feedback during posttest. Unlike generalization, posttest problems contained no perceptual support (e.g., fruit and vegetable icons and colored zeros).

To test the effect of alignment on transfer from training to generalization and posttest problems, participants were randomly assigned to one of three between-subjects experimental conditions: no alignment, multiple exemplars, and progressive alignment (cf. Figures 4a to 4c). In the no alignment condition, participants received training and generalization problems one at a time; thus, participants were not able to compare problems on which they received feedback during training (e.g., placement of 15 cherries on a 0–100 cherry line) with problems they completed during the generalization phase (e.g., 1,500 on a 0–10,000 number line). In the multiple exemplars and progressive alignment conditions, participants received the same training problems and corrective feedback as in the no alignment condition, but generalization problems were presented alongside previously solved training problems, thereby allowing children in these groups to compare generalization problems to training problems. The multiple exemplars and progressive alignment conditions differed from one another in that participants in the progressive alignment group were *only* able to compare one training–generalization pair at a time, thereby focusing on the critical similarity (e.g., 15 cherries:100 cherries::1,500:10,000). In contrast, participants in the multiple exemplars condition saw *all* training and generalization problems on one page. Participants in the multiple exemplars condition were able to compare training problems to generalization problems just like participants in the progressive alignment condition (and thus draw the relevant analogy), but participants in the multiple exemplars condition were also given the opportunity to make less useful comparisons, such as comparing training problems to training problems (cherries vs. carrots) and comparing generalization problems to generalization problems (10,000 vs. 100,000). To put it another way, participants in the progressive alignment condition were constrained in the types of comparisons they could make, whereas participants in the multiple exemplars condition were unconstrained in the comparisons they could make.

Results and Discussion

In the following sections, we will analyze children's number line estimates during training, generalization, and at posttest to determine whether children were capable of using an analogy to scale up their linear representation from 0–100 number lines to larger numerical magnitude contexts.

Training. We first examined accuracy of numerical estimates during training to confirm that all three experimental groups performed well on 0–100 number lines, which were problems from which the groups were expected to generalize in the generalization and posttest phases. As expected, a 3 (condition: no alignment, multiple exemplars, progressive alignment) × 4 (unit type: none, pear, cherry, carrot) ANOVA revealed no effect of condition or Condition × Unit Type interaction. Presumably due to the fixed order of items, however, there was a small but reliable effect of unit type, $F(3, 129) = 15.13$, $p < .001$, $\eta^2 = .24$, with less accurate estimates for the first type of units presented (none, $M = 89\%$) than for the other three types of units (pear: $M = 94\%$, $d = 0.82$; cherry: $M = 94\%$, $d = 1.21$; carrot: $M = 93\%$, $d = 1.09$; ps < .01), which did not differ significantly.

To confirm that these high levels of accuracy were associated with use of linear representations, we again compared the fit of the best fitting linear and logarithmic regression functions to median numerical estimates across three experimental conditions. Across all training problems, participants in the no alignment condition produced more linear than logarithmic series of estimates (blank: linear $R^2 = .99$ > logarithmic $R^2 = .85$; pear: linear $R^2 = .99$ > logarithmic $R^2 = .86$; cherry: linear $R^2 = 1$ > logarithmic $R^2 = .85$; carrot: linear $R^2 = 1$ > logarithmic $R^2 = .85$), as did participants in the multiple exemplars condition (blank: linear $R^2 = .99$ > logarithmic $R^2 = .86$; pear: linear $R^2 = 1$ > logarithmic $R^2 = .84$; cherry: linear $R^2 = 1$ > logarithmic $R^2 = .84$; carrot: linear $R^2 = 1$ > logarithmic $R^2 = .84$), and the progressive alignment condition (blank: linear $R^2 = .873$ > logarithmic $R^2 = .869$; pear: linear $R^2 = 1$ > logarithmic $R^2 = .84$; cherry: linear $R^2 = 1$ > logarithmic $R^2 = .83$; carrot: linear $R^2 = 1$ > logarithmic $R^2 = .83$). A similar pattern also emerged from the examination of individual children's estimates. Regardless of units, most children's estimates were best fit by the linear function in the no alignment condition ($M = 93\%$ best fit by linear), the multiple exemplars condition ($M = 88\%$ best fit by linear), and the progressive alignment condition ($M = 86\%$ best fit by linear). Further, tests of Fisher's exact probabilities revealed no association between experimental group and best fitting regression function (logarithmic versus linear; ps > .05).

Thus, all three experimental groups were likely to use an accurate, linear representation of numerical magnitudes in the 0–100 scale, suggesting they were equally likely to benefit from applying their linear representations of the 0–100 scale to 0–1,000, 0–10,000, and 0–100,000 scales. In the next section, we explored whether all three groups did in fact apply their linear representations to larger scales and whether that application was fostered by alignment of small and large numerical scales.

Generalization. To test the effect of progressive alignment on generalization of linear representations to larger scales, we first examined accuracy across three experimental groups. To do this, we conducted a 3 (condition: no alignment, multiple exemplars, progressive alignment) × 4 (scale: 0–100, 0–1,000, 0–10,000, 0–100,000) repeated measures ANOVA on accuracy scores on generalization problems. As expected, there was a main effect of scale, $F(3, 41) = 28.73$, $p < .001$, $\eta^2 = .02$, and a main effect of condition, $F(2, 43) = 10.44$, $p < .001$, $\eta^2 = .91$, although these main effects were qualified by a significant Condition × Scale interaction, $F(6, 84) = 2.99$, $p < .01$, $\eta^2 = .04$.

Further, to investigate significant condition differences in levels of accuracy across generalization problems, we conducted post hoc ANOVAs to determine how differing amounts of alignment (e.g., no alignment, multiple exemplars, progressive alignment) affected accuracy of children's estimates (see Table 2). As anticipated, there were no differences in accuracy across experimental conditions for the 0–100 blank number line problems, presumably because second graders typically represented magnitudes of these numbers as increasing linearly. However, children assigned to the multiple exemplars and progressive alignment conditions produce more accurate series of number line estimates for 0–1,000, 0–10,000, and 0–100,000 generalization problems as compared to those children assigned to the no alignment condition.

To assess whether these changes in accuracy were caused by the logarithmic-to-linear shift, we next compared the fit of the best-fitting linear and logarithmic functions to median numerical estimates across three experimental groups (Figure 5). As in performance on 0–100 problems during training, the best fitting function for all experimental conditions on the 0–100 generalization problems was the linear rather than logarithmic regression function (no alignment, linear $R^2 = 1.0$ > logarithmic $R^2 = .86$; multiple exemplars, linear $R^2 = .99$ > logarithmic $R^2 = .85$; progressive alignment, linear $R^2 = .99$ > logarithmic $R^2 = .85$).

Experimental groups differed, however, for larger generalization problems (0–1,000, 0–10,000, and 0–100,000). For the no alignment group, the best fitting function was logarithmic across the 0–1,000 and 0–10,000 generalization problems (logarithmic $R^2 = .9$ and .85, respectively); on 0–100,000 generalization problems, both the linear ($R^2 = .46$) and logarithmic function ($R^2 = .39$) provided uncommonly poor fits to children's estimates. For the multiple exemplars group, the best fitting function was the linear one for both 0–1,000 and 0–10,000 generalization problems (linear $R^2 = .96$ and .84, respectively), but the logarithmic function was the best fitting function across 0–100,000 generalization problems (logarithmic $R^2 = .86$). Finally, for the progressive alignment group, the linear function provided the best fit across 0–1,000, 0–10,000, and 0–100,000 generalization problems (linear $R^2 = 1.00$, .98, and .98, respectively).

To ensure that these condition differences in model fit were due to the effect of alignment and not due to irregularities arising from averaging, we used Fisher exact probability tests to examine reliability of association between condition and percentage of individual children best fit by the linear function. On the two smaller scales, 0–100 and 0–1,000, we observed no significant associations between condition and fit by the linear function (on

Table 2
Accuracy of Numerical Estimates by Scale and Condition

	0–1,000	0–10,000	0–100,000
Condition	$F(2, 45) = 8.38$, $p < .001$ Progressive, M = 90% > No, M = 78%, d = 1.25 Multiple exemplars, M = 89% > No, M = 78%, d = 1.09	$F(2, 45) = 6.06$, $p < .01$ Progressive, M = 83% > No, M = 63%, d = 0.95 Multiple exemplars, M = 84% > No, M = 63%, d = 1.16	$F(2, 45) = 10.89$, $p < .001$ Progressive, M = 85% > No, M = 63%, d = 1.48 Multiple exemplars, M = 82% > No, M = 63%, d = 1.43

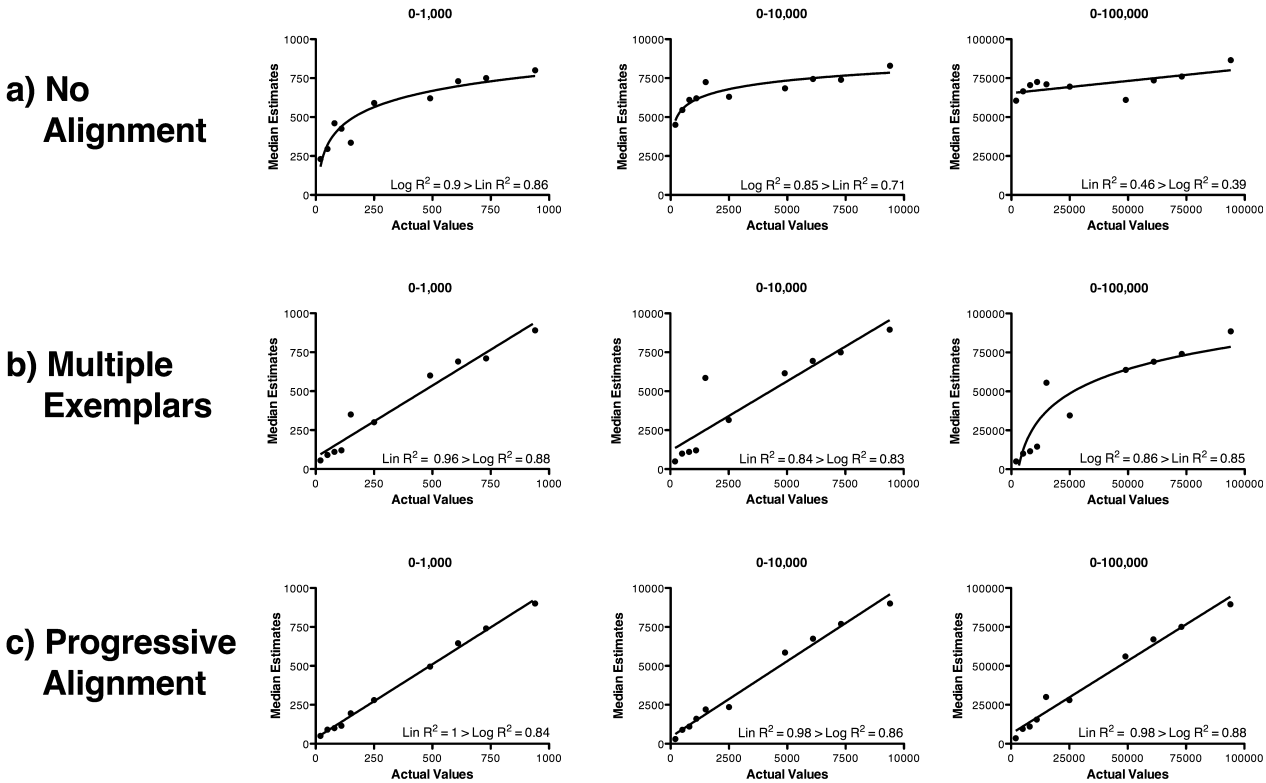

Figure 5. Experiment 2 (generalization): Median estimates of participants in the no alignment, multiple exemplars, and progressive alignment groups on 0–1,000, 0–10,000, and 0–100,000 number lines.

0–100 scales: no alignment, 93%; multiple exemplars, 100%; progressive alignment, 100%; on 0–1,000 scales: no alignment, 57%; multiple exemplars, 69%; progressive alignment, 81%). On the two larger scales, 0–10,000 and 0–100,000, however, we observed significant associations between condition and fit by the linear function, with the progressive alignment group outperforming the other two experimental groups (on 0–10,000 scales: no alignment, 50%; multiple exemplars, 56%; progressive alignment, 100%; on 0–100,000 scales: no alignment, 50%; multiple exemplars, 50%; progressive alignment, 100%; *p*s $< .05$; compare to results of Experiment, 1 illustrated in Figure 3). Thus, while children generalized linear representations of numerical magnitude to 0–1,000 scales with and without the help of alignment of small and large numerical scales, children's generalization of linear representations to 0–10,000 and 0–100,000 scales only occurred when these scales were individually aligned with 0–100 scales, as they were for children in the progressive alignment group.

In summary, findings from the generalization phase of Experiment 2—greater estimation accuracy on number lines with larger numeric values, better fits of the linear regression function to estimates on number lines with larger numeric values, and strong associations between alignment and likelihood of providing linear estimates on number lines with larger numeric values—indicate that progressive alignment of large and small number lines led children to scale up their linear representations of numerical magnitude more so than no alignment of large and small scales (no alignment condition) or alignment of units as well as alignment of large and small scales (multiple exemplars condition).

We speculate the reason children did not scale up their linear representation in the multiple exemplars condition, which presumably offered a wealth of alignment possibilities, may have resulted from learners failing to align the most effective pairs (e.g., 0–100 pears training problem compared to 0–1,000 generalization problem with one green zero). That is, in the multiple exemplars condition, children were not constrained in the types of comparisons they made across problem types because they saw all problems on one page. When faced with the ability to compare any of the number lines on the page, learners may have failed to notice surface similarities across their initial comparisons.

If the children failed to see how the problems encompassing small and large numeric scales related to one another, they may have given up on their attempts to find commonalities. Alternatively, children may have used a more shallow comparison process by simply scanning the multiple exemplars for obvious, surface commonalities as compared to engaging in a deeper comparison process that might highlight structural similarities (e.g., decimal system) across the small and large numerical scales. In the next section, we examined whether advantages of progressive alignment over other training procedures persisted to a posttest, where alignment of large and small scales desisted.

Posttest. To determine whether participants had transferred their linear representation of number to larger numbers generally or just to number lines with perceptual support (e.g., fruit and vegetable icons and colored zeros), we examined performance on a blank 0–1,000 number line that was unaccompanied by previously solved problems and without units and zeroes sharing a common color. Our reasoning was that if children had grasped the analogy we intended, they would distance their hatch marks as a linear function of actual value on number lines identical to those used in previous studies (i.e., without perceptual support).

As illustrated in Figure 6, estimates of participants in the progressive alignment group were indeed better fit by the linear function (R^2 = .95) than by the logarithmic (R^2 = .76). In contrast, children's estimates in the no alignment group were better fit by the logarithmic function (R^2 = .91) than by the linear (R^2 = .66), and children's estimates in the multiple exemplars group were also better fit by the logarithmic function (R^2 = .83) than by the linear (R^2 = .58). Finally, the same advantage of the progressive alignment group was evident in the percentage of children best fit by the linear as opposed to the logarithmic function (compare with 0–1,000 condition in Figure 3), with 63% of children in the progressive alignment group generating linear series of estimates versus 29% of children in the no alignment group, $\chi^2(1) = 3.45$, $p = .06$, and 25% of children in the multiple exemplars group, $\chi^2(1) = 4.57$, $p < .05$.

General Discussion

The ability to carry out effortless structural alignment is a hallmark of human cognitive processing (Gentner & Markman, 1997; although see Kurtz et al., 2001, for an instance of more effortful alignment) and a potentially general mechanism of cognitive change (Chen & Klahr, 1999; Gentner et al., 2001; Holyoak & Thagard, 1995; Opfer & Siegler, 2004). In our study, we investigated whether structural alignment of small numerical scales and large numerical scales might lead to a representational change in the domain of number (logarithmic-to-linear shift) that has been observed across a large number of age groups (Siegler & Booth, 2004; Opfer & Thompson, 2008; Siegler & Opfer, 2003).

Our first experiment illustrated large numeric contexts in which second-, third-, and sixth-grade children produced linear versus logarithmic series of numerical magnitude estimates. Our results supported the idea that abandoning logarithmic representations of numerical magnitude is not an all-or-none process. Rather, extending linear representations of number to ever-larger numeric scales occurs gradually over time, with half or more of each group of children continuing to use logarithmic representations on one or more tasks (e.g., 50% of sixth graders on 0–100,000; 50% of third graders on 0–10,000 and 0–100,000; and 75% of second graders on 0–1,000, 0–10,000, and 0–100,000). Thus, for large numeric scales, children's use of logarithmic representations appeared unchanged by their experiences—either due to the low frequency with which they are likely to receive any information

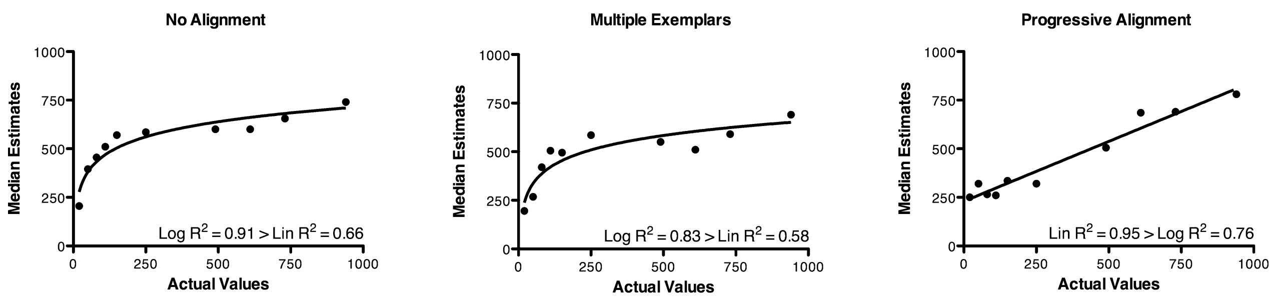

Figure 6. Experiment 2 (posttest): Median estimates of participants in the no alignment, multiple exemplars, and progressive alignment groups on unsupported 0–1,000 number lines.

about values in these large scales or due to an age-related conceptual limitation in representing large numeric values.

Our second experiment indicated that failure to abandon logarithmic representations of large numbers in favor of linear ones is likely due to poverty of input regarding large numerical values. Thus, when second graders—who produced logarithmic estimates at the highest rates (75%) across all three tasks in Experiment 1—were simply provided with the opportunity to align number line problems for small numerical scales (0–100) with number line problems of larger numerical scales (0–1,000, 0–10,000, and 0–100,000), children abandoned use of the logarithmic representation with surprising frequency. This representational change was most robust in the progressive alignment condition, where children could directly compare only relevant training and generalization problems and where 0% of children provided estimates better fit by the logarithmic than linear functions on 0–10,000 and 0–100,000 problems. Moreover, these children continued to generate linear estimates on posttest, which children in the multiple exemplars and no alignment conditions did with much less consistency. Progressive alignment of small and large numerical scales focused children's attention on the most informative surface level comparisons allowing children to search for deeper, structural commonalities as well. Children in the multiple exemplars condition were not at all constrained when attempting to compare small and large numerical scales. It is likely that surface dissimilarities (e.g., pears, cherries, carrots, 0–100, 0–1,000, 0–10,000, 0–100,000) confused learners in the multiple exemplars condition and prohibited these learners from discovering deeper, structural similarities across number lines of small and large numerical scales.

Analogy as a Mechanism of Representational Change

Why did alignment of small and large numeric scales produce rapid changes in numerical estimation and generalization of the linear representation? In our experiment, we aligned contexts in which children were familiar (e.g., 0–100) with larger, less familiar numeric contexts (e.g., 0–1,000 to 0–100,000). This alignment methodology apparently prompted second graders to scale up their linear representation of numbers by highlighting the underlying structure of the decimal system and thereby bootstrap the linear representation they already possessed in a familiar numeric context (0–100) to less familiar numeric contexts (0–1,000 through 0–100,000). Although this mechanism of representational change had been suggested in an earlier microgenetic study of numerical estimation (Opfer & Siegler, 2007), the present work provided unusually strong support for the idea by showing both that the representational change can occur without any feedback as well as by showing that the learning can be generalized very far outside the training space.

These findings on the effect of alignment in producing analogies are also apparent in a number of other empirical studies (Dixon & Dohn, 2003; Gentner, Loewenstein, & Thompson, 2003; Kotovsky & Gentner, 1996; Uttal, Schreiber, & DeLoache, 1995), where alignment has led people to compare examples and thereby extract an abstract relational schema that might be applied to other relevant problems. For example, novice negotiators who were learning about negotiation strategies were asked to compare written case studies in the hopes that a common negotiation strategy would be extracted despite differing surface features of presented problems. Novice negotiators who compared two written case studies detailing negotiation strategies were better able to abstract general problem-solving schemas which they later applied to future negotiation situations, like face-to-face negotiations, than were negotiators who were not explicitly told to compare cases or who received no case-based training whatsoever (Gentner et al., 2003; Loewenstein, Thompson, & Gentner, 2003). The general problem-solving schema was abstracted, or disembedded, from the original learning context so that it could be applied to other relevant case studies outside the original learning context.

Likewise, other investigators have similarly noted that alignment specifically and similarity comparison in general allows participants to disembed underlying structural relations in problem-solving tasks. In classic scale model problem-solving tasks conducted by DeLoache and her colleagues (e.g., DeLoache, Kolstad, & Anderson, 1991; Uttal et al., 1995), young children were shown where a hidden object could be found in a scale model and then were asked to locate the hidden object in the larger artificial room the model represented. Not only must children realize the scale model is related to the larger room and therefore locations in the model can be mapped directly onto this larger artificial room in a one-to-one fashion, but children must also hold in mind the representation of the scale model to find the hiding place of the object in the larger room (Uttal et al., 1995). According to DeLoache et al. (1991), one way that children successfully uncover the one-to-one mapping between the model and artificial room is by attending to overlapping similarity in the two spaces (e.g., object matches and size or scale of rooms). Further, Loewenstein and Gentner (2001) found that 3-year-olds' search behavior in the DeLoache hiding task was improved by their ability to compare two sequential hiding tasks transpiring across two highly similar hiding rooms whereas search behavior was not improved for those 3-year-olds who saw the highly similar hiding rooms but were not prompted to compare them.

Another example of participants' ability to abstract a general problem-solving schema through alignment is Dixon and Dohn's (2003) alternating sequence problems. In these problems, participants were shown a connected sequence of balance scales or gears and then were asked to predict the position of the last balance scale in sequence (up or down) or direction of movement of the last gear in sequence (clockwise or counterclockwise). Participants who discovered the underlying structural relation through comparison of multiple examples of alternating sequence problems often extracted the solution more quickly or readily and more consistently transferred the solution to new problem types than did those participants who were explicitly instructed on the relation uniting the problems (e.g., alternating sequences: up, down, up, down or clockwise, counterclockwise, clockwise, counterclockwise).

As in these cases of learning, children in our own experiments were given opportunities that seemed crucial for eliciting their analogies. First, they were given opportunity to compare across multiple training problems, much like novice negotiators comparing written case studies (Gentner et al., 2003), toddlers comparing scale models to larger rooms (DeLoache et al., 1991; Loewenstein & Gentner, 2001), adults comparing alternating sequence problems (Dixon & Dohn, 2003), and children comparing geometric shapes that differed in symmetry and monotonicity in match to sample tasks (Kotovsky & Gentner, 1996). Second, our participants were given perceptual matches to facilitate alignment of small number line problems (e.g., two red cherries in 0–100 problems) with large number line problems (e.g., two red zeros in 0–10,000 problems), just as DeLoache and colleagues' toddlers used high physical similarity across the size of the scale model and artificial room. Third, not only did presence of overlapping similarity lead participants to abstract structural relations, but children's superior understanding of the 0–100 scale was probably crucial, given the finding that increasing participants' knowledge base increases the likelihood of learners noticing structural relations (Kotovsky & Gentner, 1996). Thus, Experiment 2 appeared to provide many critical opportunities that children needed to draw analogies from more familiar to less familiar problems.

Alternative Role for Familiarity in Numerical Estimation: Segmented Linear Model

According to the segmented linear model, children's familiarity with numbers, as indexed by counting ability within a particular numeric scale, is a good predictor of whether children produce linear estimates (as in familiar numeric scales) or logarithmic estimates (as in unfamiliar numeric scales). Although, according to Ebersbach et al. (2008), the fit of the logarithmic function may be illusory, and children's estimates might be better fit by two linear regression functions that intersect at a ''change point.''

Results of Experiment 1 provided an unusual opportunity to evaluate this alternative model, and our results do not support the segmented linear model. Specifically, when we fit estimates using the segmented linear function, we found increasingly larger change points as numeric context increased from 0–1,000 (second grade: 158.7, third grade: 366.5, sixth grade: 386.3, adults: 421.35), 0–10,000 (second grade: 2,330, third grade: 2,500, sixth grade: 4,263, adults: 3,668), and 0–100,000 (second grade: 37,000, third grade: 26,592.5, sixth grade: 23,798, adults: 48,999.5), $r = .94$, $F(1, 11) = 70.89$, $p < .001$, suggesting that change point is a task-specific scaling parameter for a logarithmic function rather than an index of children's numeric familiarity. To test the idea that change point is just a task-specific scaling parameter, we next regressed mean age of each participant group (second grade: 7.9 years, third grade: 9.23 years, sixth grade: 12.06 years, adults: 19.96 years) against the segmented linear model's change point, and we found no correlation between age and supposed range of children's familiarity, $r = .15$, $F(1, 11) = .23$, $p > .05$, *ns*. Thus, although numeric anchor (0–1,000, 0–10,000, or 0–100,000) accounted for 94% of variation in change points, age, in comparison, accounted for an insignificant (15%) amount of variation in change points.

Although change point of the segmented linear model has been interpreted by Ebersbach et al. (2008) as reflecting a division in children's familiarity with numbers, the finding that change point varied directly with scale of the number line task (0–1,000 through 0–100,000) suggests that change point is better understood merely as a task-specific scaling parameter. This point is clearest for third graders, where data from both Ebersbach et al. (2008) and our own data indicate a change point of approximately 342 and 366 (respectively) on the 0–1,000 task. Taken alone, it is tempting to view 342–366 as an outer range where children's familiarity with numbers might end. However, since we found the same age group of children producing estimates with a change point of 2,500 and 26,592 on a 0–10,000 and 0–100,000 task (respectively) these findings clearly undercut this logic: Change point in a segmented linear model cannot simply index children's numeric familiarity. The alternative explanation—that participants' familiarity with numbers corresponds directly to the task to which they were randomly assigned—is clearly implausible. Thus, although participants in our studies appeared to use their knowledge of a familiar numeric scale (0–100) to draw an analogy to less-familiar scales (0–1,000 through 0–100,000), familiarity itself does not appear to provide a viable alternative explanation for the logarithmic pattern of estimates observed in Experiment 1.

Educational Implications

Understanding of place value, or interrelations of the numeric decimal system, is a part of children's blossoming number sense (U.S. Department of Education, 2008). This number sense, according to a recent report by the National Mathematics Advisory Panel ''includes the ability to estimate results of computations and thereby to estimate orders of magnitude'' (p. 18). The panel detailed the state of mathematics research and education in the United States and noted the importance of fostering children's number sense because ''poor number sense interferes with learning algorithms and number facts and prevents use of strategies to verify if solutions to problems are reasonable'' (p. 27).

Our current research on alignment of small and large numerical scales offers a possible means of providing students explicit as well as implicit instruction on place value and the decimal system. In Experiment 2, we showed that progressive alignment of small and large numerical scales leads young children to adopt a linear representation of number and abandon a logarithmic one, which lends support to important educational implications of our results. The educational implication of this alignment methodology is that it might be used as a simple classroom intervention that may allow children to being thinking about numbers in a more adult-like, linear fashion months and maybe even years before classical classroom instruction or life experiences would lead them to do so. As part of the simple intervention, teachers might align small numeric scales with larger ones and even provide a verbal or visual hint about how the decimal system works (e.g., crossing out the extra zero in 0–1,000 order of magnitude to show how it is similar to the 0–100 order of magnitude). To our knowledge, highlighting properties of the decimal system through alignment is not a technique commonly used by elementary school teachers, but based on previous evidence of causal and correlational links between numerical estimation and other mathematical skills (Booth & Siegler, 2006, 2008), we believe this intervention is likely to be an easy and effective one to implement.

References

Abdellatif, H. R., Cummings, R., & Maddux, C. D. (2008). Factors affecting the development of analogical reasoning in young children: A review of the literature. *Education*, *129*, 239–249.

Alexander, P. A., Willson, V. L., White, C. S., & Fuqua, J. D. (1987). Analogical reasoning in young children. *Journal of Educational Psychology*, *79*, 401–408.

Banks, W. P., & Hill, D. K. (1974). The apparent magnitude of number scaled by random production. *Journal of Experimental Psychology Monograph*, *102*, 353–376.

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology*, *41*, 189–201.

Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development*, *79*, 1016–1031.

Brannon, E. M. (2005). What animals know about numbers. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 85–107). New York: Taylor & Francis.

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition of the control of variables strategy. *Child Development*, *70*, 1098–1120.

Dehaene, S. (2007). Symbols and quantities in parietal cortex: Elements of a mathematical theory of number representation and manipulation. In P. Haggard, Y. Rossetti, & M. Kawato (Eds.), *Attention and performance XXIV. Sensorimotor foundations of higher cognition* (pp. 527–574). Cambridge, MA: Harvard University Press.

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

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition*, *43*, 1–29.

DeLoache, J. S., Kolstad, V., & Anderson, K. N. (1991). Physical similarity and young children's understanding of scale models. *Child Development*, *62*, 111–126.

Dixon, J. A., & Dohn, M. C. (2003). Redescription disembeds relations: Evidence from relational transfer and use in problem solving. *Memory & Cognition*, *31*, 1082–1093.

Doumas, L. A. A., Hummel, J. E., & Sandhofer, C. M. (2008). A theory of the discovery and predication of relational concepts. *Psychological Review*, *115*, 1–43.

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.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307–314.

Furlong, E. E., & Opfer, J. E. (2009). Cognitive constraints on how economic rewards affect cooperation. *Psychological Science*, *20*, 11–16.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, *44*, 43–74.

Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development*, *78*, 1343–1359.

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.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*, 155–170.

Gentner, D., Brem, S., Ferguson, R. W., Markman, A. B., Levidow, B. B., Wolff, P., et al. (1997). Analogical reasoning and conceptual change: A case study of Johannes Kepler. *Journal of the Learning Sciences*, *6*, 3–40.

Gentner, D., Holyoak, K. J., & Kokinov, B. N. (Eds.). (2001). *The analogical mind: Perspectives from cognitive science*. Cambridge, MA: MIT Press.

Gentner, D., Loewenstein, J., & Hung, B. (2007). Comparison facilitates children's learning of names for parts. *Journal of Cognition and Development*, *8*, 285–307.

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, *95*, 393–408.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, *52*, 45–56.

Gentner, D., & Namy, L. (1999). Comparison in the development of categories. *Cognitive Development*, *14*, 487–513.

Goswami, U., & Brown, A. (1990). Melting chocolate and melting snowman. Analogical reasoning and causal relations. *Cognition*, *35*, 69–95.

Griffin, S., Case, R., & Siegler, R. S. (1994). Rightstart: Providing the central conceptual prerequisites for first formal learning of arithmetic to students at risk for school failure. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 25–49). Cambridge, MA: MIT Press/Bradford Books.

Holyoak, K., Junn, E. N., & Billman, D. O. (1984). Development of analogical problem-solving skill. *Child Development*, *55*, 2042–2055.

Holyoak, K., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press/Bradford Books.

Hummel, J. E., & Holyoak, K. J. (2003). A symbolic connectionist theory of relational inference and generalization. *Psychological Review*, *110*, 220–264.

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development*, *67*, 2797–2822.

Kurtz, K. J., Miao, C., & Gentner, D. (2001). Learning by analogical bootstrapping. *Journal of the Learning Sciences*, *10*, 417–446.

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*, *76*, 1723–1743.

Loewenstein, J., & Gentner, D. (2001). Spatial mapping in preschoolers: Close comparisons facilitate far mappings. *Journal of Cognition and Development*, *2*, 189–219.

Loewenstein, J., Thompson, L., & Gentner, D. (2003). Analogical learning in negotiation teams: Comparing cases promotes learning and transfer. *Academy of Management Learning and Education*, *2*, 119–127.

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*, 843–849.

Opfer, J. E., & Siegler, R. S. (2004). Revisiting preschoolers' *living things* concept: A microgenetic analysis of conceptual change in basic biology. *Cognitive Psychology*, *49*, 301–332.

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

Opfer, J. E., Thompson, C. A., & Furlong, E. (2010). Early development of spatial-numeric associations: Evidence from spatial and quantitative performance of preschoolers. *Developmental Science*, *13*, 761–771.

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.

Rattermann, M. J., & Gentner, D. (1998). More eveidence for a relational shift in the development of analogy: Children's performance on a causal-mapping task. *Cognitive Development*, *13*, 453–478.

Roberts, W. A. (2005). How do pigeons represent numbers? Studies of number scale bisection. *Behavioral Processes*, *69*, 33–43.

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

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

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., & Ramani, G. B. (2008). Playing linear numerical board games promotes low-income children's numerical development. *Developmental Science, Special Issue on Mathematical Cognition*, *11*, 655–661.

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

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, & Education*, *3*, 143–150.

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.

U.S. Department of Education (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. Washington, DC: U.S. Department of Education.

Uttal, D., Schreiber, J. C., & DeLoache, J. S. (1995). Waiting to use a symbol: The effect of delay on children's use of models. *Child Development*, *66*, 1875–1889.

Xu, F., & Spelke, E. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*, B1–B11.