

Development of Numerical Estimation in Young Children

Robert S. Siegler and Julie L. Booth

Two experiments examined kindergartners', first graders', and second graders' numerical estimation, the internal representations that gave rise to the estimates, and the general hypothesis that developmental sequences within a domain tend to repeat themselves in new contexts. Development of estimation in this age range on 0-to-100 number lines followed the pattern observed previously with older children on 0-to-1,000 lines. Between kindergarten and second grade (6 and 8 years), patterns of estimates progressed from consistently logarithmic to a mixture of logarithmic and linear to a primarily linear pattern. Individual differences in number-line estimation correlated strongly with math achievement test scores, improved estimation accuracy proved attributable to increased linearity of estimates, and exposure to relevant experience tended to improve estimation accuracy.

Estimation is a pervasive process in the lives of both children and adults. How much time will it take to get home? How much money will the food in the grocery cart cost? How heavy is this object? How far is the distance between here and there? How many weeks will it take to write this paper? Without the ability to estimate reasonably accurately, life would be difficult.

Despite the importance of estimation both in the classroom and in everyday life, far less is known about its development than about the development of other basic quantitative abilities, such as subitizing, counting, and adding (Dowker, in press; Geary, 1994). One reason for the discrepancy is that estimation subsumes a much greater range of tasks and knowledge than the better understood quantitative processes. Some estimation tasks—for example, estimating distance, time, or money—require knowledge of measurement units such as miles, minutes, or dollars; other estimation tasks—for example, estimating the number of people in a room or dots on a page—do not. Similarly, some uses of estimation—for example, estimating national populations and land areas—require prior knowledge of the entities whose properties are being estimated; other

uses of estimation—for example, estimating the number of candies in a jar—do not. This variability of tasks and prerequisite knowledge has made it difficult to identify the processes that unite all types of estimation and to formulate experimental paradigms that are useful for investigating their development.

The present study is based on an explicit assumption about the core process of estimation: Estimation is a process of translation between alternative quantitative representations. Some estimates involve non-numerical-to-non-numerical translations, for example, translating perceived brightness into line lengths. Other estimates involve numerical-to-numerical translations, for example, translating a multidigit multiplication problem into an estimated product. Yet other estimates involve numerical-to-non-numerical translation, for example, presenting children with a number and asking them to locate its position on a number line.

This last task, the one presented in the current study, was appealing for several reasons. First, the task is a relatively pure measure of numerical estimation in that it does not require knowledge of measurement units or particular entities. Second, the task is ecologically valid; many classrooms, including those of the participants in the study, include number lines, and many teachers, again including those in this study, use number lines to teach numerical concepts. Third, the task makes possible relatively straightforward tests of alternative models of numerical representation (as described later).

The task also allowed us to examine a broader question about development. Classic developmental theorists, including Piaget and Inhelder (1956),

Robert S. Siegler and Julie L. Booth, Department of Psychology, Carnegie Mellon University.

Thanks are due to John Opfer, who provided insightful comments on an earlier version of the manuscript; to the children, teachers, and principals of the Penn Hills School District; and to National Institutes of Health Grant HD19011 and Office of Educational Research and Instruction Grant R305H020060, which supported the research.

Correspondence concerning this article should be addressed to either Robert S. Siegler or Julie L. Booth, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213. Electronic mail may be sent to rs7k@andrew.cmu.edu or Juliebooth@andrew.cmu.edu.

Vygotsky (1934/1962), and Werner (1957), proposed that developmental changes at different ages and over different time spans show extensive parallels. The present study examined whether this perspective can help modern investigators anticipate patterns of change with age and experience, even in domains such as numerical estimation, which none of the classic theorists studied. In the following sections we summarize existing understanding of the development of estimation and describe the current study and how it seeks to advance that understanding.

Current Understanding of the Development of Numerical Estimation

The most consistent conclusion reached by investigators of the development of estimation is that **young children are not very skillful estimators**. This conclusion has been reached by investigators studying estimation of various properties, including distance (Cohen, Weatherford, Lomenick, & Koeller, 1979), money (Sowder & Wheeler, 1989), number of discrete objects (Hecox & Hagen, 1971), and answers to arithmetic problems (LeFevre, Greenham, & Naheed, 1993). The problem has been ascribed to various causes: mindless symbol manipulation, reliance on procedures rather than principles, lack of number sense, and lack of relevant central conceptual structures (Case & Sowder, 1990; Hiebert & Wearne, 1986; Joram, Subrahmanyam, & Gelman, 1998; Sowder, 1992). Another factor that may play a large role is reliance on inappropriate representation of numbers.

Alternative representations of numbers. Several groups of investigators have hypothesized that children's estimation reflects their internal representation of numbers. However, the hypothesized representations vary considerably. Dehaene (1997) suggested that people of all ages from infancy to adulthood, as well as many other animals, rely on a logarithmic ruler representation. Relative to a linear representation of numbers, a logarithmic representation exaggerates the distance between the magnitudes of numbers at the low end of the range and minimizes the distance between magnitudes of numbers in the middle and upper ends of the range. Thus, within a logarithmic representation such as that in Figure 1, the psychological distance between the numbers 1 and 75 is greater than that between 75 and 1,000. Reliance on the logarithmic ruler representation is said to "occur as a reflex" (Dehaene, 1997, p. 78), one that cannot be inhibited. Dehaene presented considerable evidence consistent with the use of this representation. For example, when adults

are asked to draw lines of whatever length they wish or to generate numbers randomly, the line lengths and numbers fit a logarithmic function; both adults' and children's solution times on magnitude-comparison problems fit the same function (Banks & Hill, 1974; Sekuler & Mierkiewicz, 1977).

Gibbon and Church (1981) proposed an alternative representation, which they labeled the accumulator model. Like the logarithmic ruler model, the accumulator model has been hypothesized to be used by people of all ages (Brannon, Wusthoff, Gallistel, & Gibbon, 2001). The basic claim is that numbers and other quantities are represented as equally spaced, linearly increasing magnitudes with scalar variability. This last property involves representations becoming noisier, and therefore more variable, with increasing magnitude; amount of variability increases linearly with the number being represented. Commenting on the possibility that people might be capable of other representations, in particular, the logarithmic ruler representation, Brannon et al. (2001, p. 243) argued, "Given the current state of knowledge, we view the idea that number is represented both linearly and logarithmically as unparsimonious." Like Dehaene (1997), proponents of the accumulator model have presented a considerable body of evidence consistent with the proposed

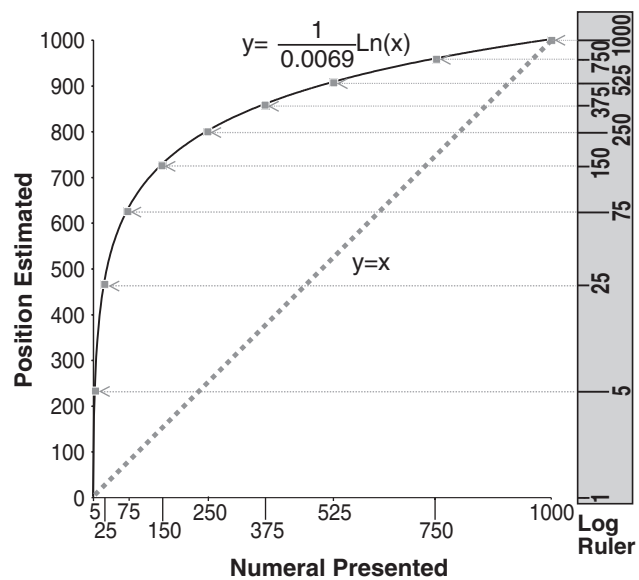


Figure 1. Predicted estimates of the logarithmic ruler and linear ruler models, when both are constrained to pass through the endpoints of the number line. Relative to the linear function, the logarithmic ruler model exaggerates distances at the low end of the numerical scale and understates them at the high end. Thus, if a child relies on a logarithmic representation, the psychological distance between 0 and 75 would be greater than that between 75 and 1,000.

representation. For example, Huntley-Fenner (2001) found that on a dot-estimation task, 5- to 7-year-olds' means and variances fit the predictions of the model.

Case and Okamoto (1996) proposed a third hypothesis, one that depicted children of different ages as using different representations but children of a given age as using a single representation. Of particular relevance to the present study, they proposed that 4- and 5-year-olds possess only a qualitative central conceptual structure for representing numbers (e.g., "This collection has a little and this one a lot"), whereas 6-year-olds and older children possess and consistently rely on a linear structure. Before acquisition of the linear structure, which Case depicted as a number line with linearly increasing magnitudes and which we refer to as the linear ruler representation, accurate numerical estimation is said to be impossible (Case & Sowder, 1990). Like the proponents of the logarithmic ruler and accumulator models, Case and his colleagues presented a great deal of evidence consistent with use of the hypothesized representation. For example, they showed that 4-year-olds generally could not accurately estimate which of two single-digit numbers was closer to a third, whereas 6-year-olds generally could (Case & Okamoto, 1996).

Multiple representations hypothesis. All three of these models provide plausible accounts of how people represent numbers and other quantities. However, none of the models seems likely to be people's only representation. Instead, it seems probable that individuals know and use multiple representations of numbers, that contextual variables influence which representation is chosen in a given situation, and that the range of situations in which children rely on each representation changes with age and numerical experience.

These hypotheses reflect not only the particulars of estimation but also a broader approach, overlapping waves theory (Siegler, 1996). In a wide range of domains, people know and use multiple rules, strategies, and representations. It seems likely that this is the case with estimation as well. In particular, from infancy onward, children may be capable of using logarithmic and accumulator representations. Subsequent experience with the formal number system in counting, arithmetic, and other numerical contexts may lead children to add linear representations, as well as a variety of categorical representations of numbers (odd-even, square-nonsquare, decade name-other, etc.) to the earlier representations. Rather than any one representation being *the* representation of numbers, varied representations may coexist and compete, with different representations being used

most often in different situations. With age and experience, children may rely increasingly on the most appropriate representation for the situation.

The use of multiple representations of numerical magnitude makes sense because different representations are most advantageous under different circumstances and at different points in the learning process. Consider logarithmic and linear representations, the representations that figure most prominently in the present study. Logarithmic representations seem especially useful as an initial representation of an unfamiliar range of numbers because such representations discriminate to a greater extent among numbers at the low end of the range than do linear representations. When children are first learning about the numbers 0 to 100, such enhanced discrimination among the magnitudes of numbers at the low end of the range is especially useful because these numbers come into play more often in single-digit addition and subtraction, counting, and other early numerical activities than do numbers higher in the range (Ashcraft & Christy, 1995; Hamman & Ashcraft, 1986). On the other hand, linear representations discriminate more clearly in the middle and high end of the range (Figure 1). Discriminating among the magnitudes of these larger numbers becomes increasingly important as children encounter, and try to understand the results of, multidigit addition and subtraction and single-digit multiplication and division problems in first and second grades.

Support for the multiple-representations perspective, and against the view that people rely on any single representation of numbers, was provided by Siegler and Opfer (2003). They asked second-, fourth-, and sixth-grade students and adults to estimate the placement of numbers on number lines. Some of the lines had endpoints of 0 and 100, whereas other lines had endpoints of 0 and 1,000; the lines were otherwise unmarked.

A striking developmental change, from reliance on a logarithmic representation to reliance on a linear representation, occurred between second and sixth grade on the 0-to-1,000 lines. The logarithmic function accounted for 95% of the variance in second graders' median estimates on these number lines, whereas the best fitting linear function accounted for only 63%. In contrast, sixth graders' and adults' estimates fit a perfectly linear pattern; the best fitting linear function accounted for 100% of the variance in median estimates at each age, whereas the logarithmic function accounted for 73% and 78% of variance among sixth graders and adults respectively. Analyses of individual participants' estimates showed

the same developmental sequence. For example, the estimation patterns of 91% of second graders were better fit by a logarithmic function than by the best fitting linear function; the corresponding figure for adults was 0%.

Findings on the 0-to-100 task indicated that the logarithmic representation was not the only one that second graders could use. Almost half (43%) generated logarithmic patterns of estimates in the 0-to-1,000 context but linear patterns in the 0-to-100 context. The task that children were asked to perform also exerted an influence; although only 9% of second graders generated a linear pattern on the 0-to-1,000 number line when presented a number and asked to mark its position with a hatch mark, 50% of the same children did so when presented a hatch mark on the line and asked to indicate the number that corresponded to it. Thus, estimation patterns, and presumably the representations used in generating them, vary with age, numerical context, and task.

Siegler and Opfer's (2003) results argued against the logarithmic ruler, accumulator, and linear ruler models of estimation and indeed against any model that hypothesizes use of a single representation. Arguing against the logarithmic ruler model was the finding that all adults and sixth graders generated a linear pattern of estimates on both 0-to-100 and 0-to-1,000 number lines. Arguing against both the accumulator model and the linear ruler model was the finding that almost all second graders and about half of the fourth graders generated a logarithmic pattern of estimates on the 0-to-1,000 number line. Arguing against all three models and any other model that proposes that people of a given age, or people in general, possess only a single representation of numbers, were the effects of task and numerical context within individual children.

The Present Study

Siegler and Opfer's (2003) findings raised several questions about the development of estimation. One question concerned parallels in developmental sequences at different ages in different numerical contexts. On the number-to-position 0-to-1,000 task, second graders consistently relied on the logarithmic representation, fourth graders sometimes did the same and sometimes relied on the linear representation, and sixth graders consistently relied on the linear representation. The ideas of the classic theorists suggested that this sequence on the 0-to-1,000 number lines might have been preceded by a parallel sequence on 0-to-100 lines. The change was expected to occur between kindergarten and second

grade for this numerical range because it is in this period that most children first gain extensive experience with this range of numbers.

A second purpose of the study was to test the relation of number-line estimation to broad measures of mathematical understanding, in particular, math achievement test scores. Case and Okamoto (1996) proposed that construction of the linear ruler representation of numbers at around age 6 allows children to solve a wide range of mathematical problems that they could not solve previously. These investigators' arguments for the central importance of such a representation in early mathematics learning suggest that individual differences in number line estimation among 5- to 7-year-olds would be positively related to individual differences in math achievement test scores. On the other hand, Dowker's (in press) review of studies of estimation indicated that even within the category of estimation tasks, individual differences on one task are frequently unrelated or minimally related to individual differences on others. The present research allowed examination of whether number-line estimation was in fact related to math achievement.

A third purpose of the study was to test the contributions of two potential sources of age-related improvement in estimation: increasing reliance on linear representations and increasing precision of estimates. Siegler and Opfer's (2003) explanation for why estimates became increasingly accurate with age was that children increasingly relied on linear representations. Another (nonexclusive) possibility, however, was that estimates might become increasingly accurate because of age-related improvements in ability to place estimates in their intended positions. Supporting this view was the finding that the absolute fit of the best fitting function to individual children's estimates tended to increase with age and experience. This possible source of age-related improvement seemed even more likely to be influential in the current study of 5- to 7-year-olds because of young children's imperfect motor control and generally high variability of thinking. To assess this influence on estimates, we asked children to estimate the position of each number twice (separated by an average of 24 other estimates). The difference between the two estimates of the same number provided a measure of the variability of estimates, independent of the linearity of the medians of the pairs of estimates.

A fourth question concerned the malleability of number-line estimation. If children were presented relevant experience, would their estimation on this task rapidly improve? If so, children are capable of

generating more advanced representations of number than they typically display. This hypothesis was tested in Experiment 2 and is described in more depth in that context.

In Experiment 1, we presented kindergartners, first graders, and second graders with 48 number-line-estimation items, 2 each of 24 numbers between 0 and 100. The central prediction was that there would be a developmental progression from predominantly logarithmic to a mix of logarithmic and linear to predominantly linear patterns of estimates. Other predictions were that estimation accuracy would correlate positively with math achievement test scores; that both increasing linearity and decreasing variability would contribute to age-related improvements in estimation accuracy; and that linearity, variability, and accuracy of estimates all would improve with age and grade.

Experiment 1

Method

Participants. Participating in this study were 85 students (44 males, 41 females): 21 kindergartners (mean age = 5.8, $SD = .33$), 33 first graders (mean age = 6.9, $SD = .36$), and 31 second graders (mean age = 7.8, $SD = .31$). Among the participants, 67% were Caucasian, 32% were African American, and 1% were Asian. One child did not seem to understand the task and therefore was excluded from consideration. The experimenter was a Caucasian female graduate student, the second author.

Children were recruited from three schools that included both middle- and low-income families. The percentages of children in the three schools who were eligible for the free or reduced-fee lunch program were 41%, 43%, and 49%. Teacher responses to a questionnaire indicated that number lines were used as part of the curriculum in all classrooms. Participation in the experiment was voluntary, and children received no tangible compensation for taking part.

Materials. Stimuli for the experiment were 48 sheets of paper, each with a 23-cm line printed across the middle, with 0 at the left end and 100 at the right end. A number between 0 and 100 was printed at the top of each page. To improve our ability to discriminate between linear and logarithmic estimation patterns, numbers below 30 were oversampled, with 10 numbers between 0 and 30 and 14 numbers between 30 and 100. The 24 numbers presented were 3, 4, 6, 8, 12, 17, 21, 23, 25, 29, 33, 39, 43, 48, 52, 57, 61, 64,

72, 79, 81, 84, 90, and 96. Within each set of 24 number line sheets, the pages were ordered randomly.

Each participant's Stanford Achievement Test score (SAT-9) for mathematics was also obtained. The test was administered to all children in the district near the end of the academic year, 3 to 4 months after the experimental session. The SAT-9 is a multiple-choice academic achievement test for which national norms are available (Assessment Resource Center, 2002).

Procedure. Children met one on one with the experimenter for a 20-min session at a time deemed appropriate by their teacher. At the beginning of the session, children were told that they would be given number lines with 0 at one end and 100 at the other end and that they would show the experimenter where they thought different numbers would fall on the line by marking the right location with a pencil. No feedback was provided about any of their marks, though the experimenter periodically offered general praise for doing a good job.

Results

Accuracy of estimates. To obtain an overall sense of the accuracy of children's estimates, we computed each child's percent absolute error:

$$\left| \frac{\text{Estimate} - \text{Estimated Quantity}}{\text{Scale of Estimates}} \right|$$

To illustrate how this measure works, if a child was asked to estimate the location of 72 on a 0-to-100 number line and placed the mark at the point that corresponded to 84, the absolute error would be 12%, $(84 - 72)/100$.

A one-way analysis of variance (ANOVA) on the relation of grade to percent absolute error indicated that individual kindergartners' estimates were considerably farther from the correct position than were those of first or second graders (percent absolute error = 27%, 18%, and 15%, respectively), $F(2, 82) = 21.65$, $p < .01$, $\eta^2 = .35$. These results indicated that the accuracy of estimation in the current study was comparable to that obtained by Siegler and Opfer (2003) for second graders, the one age group included in both experiments (12% in the earlier study, 15% here with a lower socioeconomic status sample).

Pattern of estimates. The initial analyses in this set examined the fit of linear, logarithmic, and exponential functions to the median estimates of children, grouped by grade level, for each of the 24 numbers that were presented. Medians rather than means were used to minimize the effect of outliers.

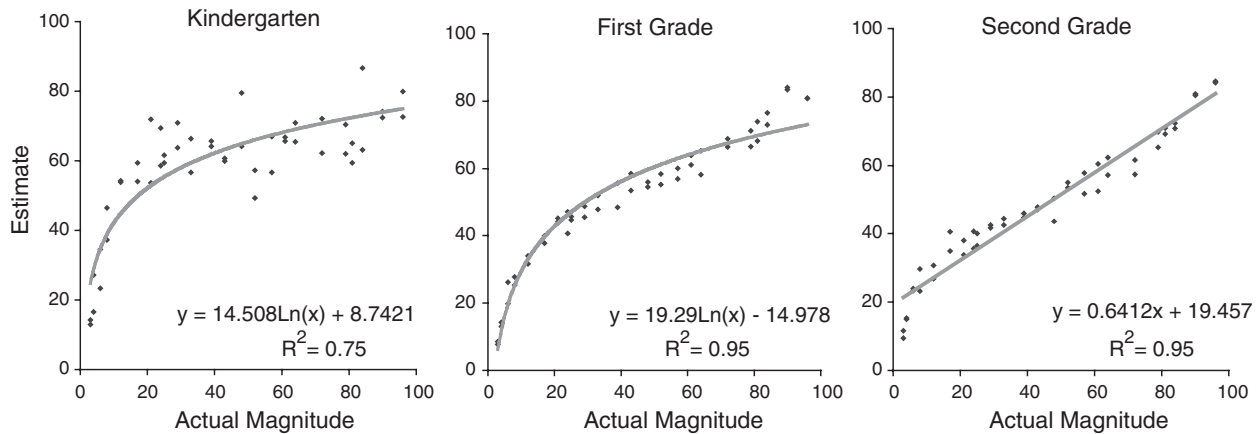


Figure 2. Progression from logarithmic pattern of median estimates among kindergartners (left panel) to linear pattern of estimates among second graders (right panel) in Experiment.

The exponential function fit the data much less well than the other two functions; therefore, it was not included in further group-level comparisons.

As shown in Figure 2, the kindergartners' median estimates for each number were better fit by the logarithmic function than by the linear function ($R^2 = .75$ vs. $.49$). To test the significance of the difference, we performed a paired-sample t test comparing the absolute value of the difference between the median of children's estimates for each number and the prediction for that number generated by: (a) the best fitting linear model and (b) the logarithmic model. The difference between the predictions of the logarithmic function and the kindergartners' median estimates was significantly less, indicating that the logarithmic function better fit the kindergartners' estimates, $t(47) = 3.84$, $p < .01$, $d = .55$. In contrast, second graders' median estimates were better fit by the linear function than by the logarithmic function ($R^2 = .95$ vs. $.88$), $t(47) = 2.99$, $p < .01$, $d = .43$. At the first-grade level, the logarithmic function and the linear function fit about equally well ($R^2 = .95$ and $.90$), $t(47) = 1.69$, $p = .10$, $d = .24$. (The fact that the absolute fit was high for both functions reflects the mathematical relation between the two; for perfectly linear data with a slope of 1, a logarithmic function that goes through both endpoints accounts for 84% of the variance.)

To test whether these findings regarding group medians also fit individual children's performance, we examined the fit of linear, logarithmic, and exponential functions to each child's estimates. The model that fit the most individual children varied with age, $\chi^2(2, N = 80) = 12.33$, $p < .01$. The percentage of children whose estimates were best fit by the logarithmic function decreased from 81% to 64% to

45% between kindergarten and second grade. The percentage of children whose estimates were best fit by the linear function increased from 5% to 30% to 55% over the same period. Performance of the remaining 14% of kindergartners and 6% of first graders was best fit by the exponential function.

Next, to quantify changes with age in the linearity of individual children's estimates, the variance accounted for by the best fitting linear function for each child's estimates was examined. The best fitting linear function accounted for an average of 24% of the variance in individual kindergartners' estimates, 59% for first graders, and 64% for second graders. First and second graders' estimates fit the linear function considerably better than did kindergartners', $F(2, 82) = 23.63$, $p < .01$, $\eta^2 = .37$.

In addition to estimates becoming more linear with increasing age and experience, the slopes of the best fitting linear function also moved increasingly toward 1.00, the ideal slope relating estimates to numbers presented. The mean slope of individual kindergartners' estimates was lower than the slopes of first and second graders (mean slopes = $.33$, $.58$, and $.60$, respectively), $F(2, 82) = 11.23$, $p < .01$, $\eta^2 = .22$.

Variability of estimates. To examine age or grade trends in the variability of estimates, we conducted a one-way ANOVA on the relation of grade to the mean for each child of the absolute difference between the two estimates for each number. As anticipated, the mean difference between individual children's two estimates decreased from kindergarten (17.5) to first and second grades (11.6 and 10.7, respectively), $F(2, 82) = 14.43$, $p < .01$, $\eta^2 = .26$. Contrary to the prediction of the accumulator model, but consistent with Siegler and Opfer's (2003) findings

with older children, variability of estimates was unrelated to the magnitude of the numbers being estimated. Numerical magnitude accounted for 0%, 2%, and 7% of the variance in variability of estimates for kindergartners, first graders, and second graders, respectively.

Linearity, variability, and accuracy of estimates. The parallel age trends in linearity, variability, and accuracy suggested that the improvements in linearity and variability might both contribute to improvements in the accuracy of estimates. To test whether this was the case, we computed a pair of regression analyses of the separate contributions of each variable to estimation accuracy. One regression analysis involved entering one variable before the other and observing the additional variance contributed by the second variable. The other regression analysis involved entering the two variables in the opposite order.

As illustrated in Table 1, degree of linearity proved to be the sole unique predictor of estimation accuracy. Regardless of whether the analysis was computed across grades or within grade, the linearity of each child's estimates invariably added at least 20% of variance in the child's percent absolute error to that accounted for by the variability of the child's estimates. In contrast, the variability of each child's estimates never added more than 1% to the variance accounted for by linearity of the child's estimates. Thus, despite the parallel age trends of linearity, variability, and accuracy, the decreases in variability of estimates did not contribute to the increases in accuracy with age and grade.

Number-line estimation and math achievement. To examine the relation of the quality of number-line estimation to a broad measure of mathematical proficiency, we computed partial correlations, controlling for age within grade, between each child's estimation accuracy and that child's performance on

the math section of the SAT-9. Percent absolute error predicted math achievement test scores at all three grade levels: kindergarten, $pr(18) = -.45$, $p < .05$; first grade, $pr(30) = -.66$, $p < .01$; and second grade, $pr(28) = -.37$, $p < .05$. Within each grade, the smaller a child's percent absolute error of estimates, the higher was that child's achievement test score.

To test whether this result was unique to the measure of estimation quality, we performed parallel analyses using an alternative measure of quality: variance accounted for by the linear function. The results paralleled those found previously. The fit of the linear function to each child's estimates correlated with the child's math achievement test score at all three grade levels: kindergarten, $pr(18) = .57$, $p < .01$; first grade, $pr(28) = .69$, $p < .01$; and second grade, $pr(28) = .39$, $p < .05$. The more linear a child's number line estimates, the higher was the child's achievement test score.

Discussion

The results provided strong support for the hypotheses that motivated the experiment. The same developmental sequence—from predominant reliance on a logarithmic representation to a mixture of reliance on logarithmic and linear representations to predominant reliance on a linear representation—was present among kindergartners through second graders on the 0-to-100 number lines as had been present among second through sixth graders on the 0-to-1,000 lines. Accuracy of estimation on the number-line task correlated with math achievement at all three grade levels. The only departure from expectation was that increasing linearity of estimates appeared to be the sole source of the improvement in estimation accuracy. Variability of estimates showed the expected decrease with age and experience, but it

Table 1
Linearity and Variability as Predictors of Percent Absolute Error: Experiment 1

Analysis	Sample	R^2_{Lin}	$R^2_{\text{Lin+Var}}$	Added R^2	Significance
1. Linearity entered before variability	All children	.79	.79	0	<i>ns</i>
	Kindergartners	.21	.21	0	<i>ns</i>
	First graders	.79	.80	.01	<i>ns</i>
	Second graders	.76	.77	.01	<i>ns</i>
2. Variability entered before linearity	All children	R^2_{Var}	$R^2_{\text{Var+Lin}}$		
	Kindergartners	.21	.79	.58	$p < .01$
	First graders	.01	.21	.20	$p < .05$
	Second graders	.10	.80	.70	$p < .01$
		.13	.77	.64	$p < .01$

did not explain any variance beyond that which could be explained by the linearity of estimates.

The results of Experiment 1 also raised the question: How malleable is young children's estimation? Individual children's estimates became more accurate and linear with age and experience, but even in second grade, the estimates remained far from perfect. Second graders' estimates were off by an average of 15%, the best fitting linear function accounted for an average of 64% of the variance in their estimates, and the linear function was the best fitting for only 55% of children. The second graders' estimates were superior to those of kindergartners but well below the levels attained on the same task by sixth graders and adults in Siegler and Opfer (2003).

Experiment 2

Experiment 2 was designed to test whether the accuracy and linearity of young children's estimates could be increased by a procedure designed to trigger cognitive conflict. The experiment involved asking children first to locate 10 evenly spaced numbers on a single number line; then to think about these initial estimates and, if desired, to revise the estimates on the original number line; and then to generate final estimates of all of the numbers' locations on a new number line. This procedure was expected to increase accuracy and reliance on linear representations because children who were capable of generating linear representations but were not relying on them on this estimation task would see that numbers that should be evenly spaced were not. The main hypothesis was that children who received this experience would subsequently generate more accurate and more linear representations when presented the original task involving a single estimate on each number line than children in a control group who were given additional trials on the original one-estimate-per-number-line task.

A second change in the procedure, presentation of an orienting trial at the beginning of the experiment, was adopted to insure that children understood the number-line task. On the orienting trial, all children first were asked to estimate the location of a number on a number line, then received feedback on the number's correct location, then were shown the location of their estimate alongside the correct location, and then were told why the correct location was correct. The logic was that this experience might help children understand the task.

Experiment 2 included four phases. In the first phase, children received the orienting trial. In the second phase, they received a pretest identical to the

Experiment 1 number-line estimation task. In the third phase, they received the experimental or control manipulation. In the fourth phase, they received a posttest much like the pretest. The main hypotheses were that the experimental manipulation would increase the accuracy and linearity of estimates and that the other results would replicate those from Experiment 1.

Method

Participants. Participating in Experiment 2 were 60 children (27 males, 33 females): 20 kindergartners (mean age = 6.1, $SD = .31$), 19 first graders (mean age = 7.1, $SD = .44$), and 21 second graders (mean age = 8.2, $SD = .71$). All of the children were recruited from three schools in the same school district as in Experiment 1. Participation was voluntary, and students received no compensation for taking part in the study. The experimenter was the same Caucasian female graduate student as in Experiment 1.

The populations of the schools were similar to those of the schools whose students participated in Experiment 1, both in racial composition and in economic status. Among Experiment 2 participants, 71% were Caucasian, 27% were African American, and 2% were Asian. The percentages of children in each school who were eligible for the free or reduced-fee lunch program were 35%, 43%, and 49%. As in Experiment 1, teachers in all classrooms reported that they used number lines in their curriculum.

At the outset of the study, children in each of the three grades were randomly assigned to the experimental or control group. More than 90% of children completed all four phases of the experiment; however, 5 children did not. Two children were absent on all days when the posttest was administered, 1 child's family moved out of the district before the posttest was given, and 2 children did not understand the task (e.g., 1 put every estimate at the midpoint of the number line). Performance of the 2 children who did not understand the task was not included in any data analysis; performance of the 3 children who completed the pretest but not the posttest was only included in analyses of pretest data.

Procedure and materials. The Experiment 2 procedure was similar to that in Experiment 1 but included two sessions, each with two phases. The two sessions were separated by roughly a week, with each session taking about 20 min.

The first phase of Session 1 was an orienting trial. The single feedback problem in this phase (the only

feedback in the experiment) differed slightly for the older and younger children. First and second graders were presented a 0-to-100 number line and asked to mark the location of 50. Following the child's estimate, the experimenter presented a number line that indicated the correct placement of 50, compared the correct placement with the child's estimate, and explained that the number being estimated was halfway between the endpoints so that its location should also be halfway between them. Kindergartners were presented a number line with 10 rather than 100 at the right end and were asked to mark the location of 5 on it. The reason the kindergartners were presented the 0-to-10 number line was that their greater familiarity with the numbers 0 to 10 increased the likelihood that they would understand the lesson of the orienting task.

In the second phase of Session 1 (the pretest), all children were presented the same 24 estimation items twice each on standard 0-to-100 number lines as in Experiment 1. Kindergartners were also presented an additional 18 estimation items on 0-to-10 number lines, 2 presentations each of the numbers 1 through 9.

The procedure followed in the first phase of Session 2 depended on whether the child was in the experimental or control group, with the problems again varying with the child's age. In the control group, first and second graders were presented the same 24 numbers on the same 0-to-100 number line as on their pretest; kindergartners were presented the same 9 numbers on 0-to-10 number lines as on the 0-to-10 portion of their pretest.

In the experimental group, first and second graders were presented a single number line with 10 numbers printed in random order above it (5, 15, 25, 35, 45, 55, 65, 75, 85, and 95); kindergartners were presented the numbers 1 to 10 inclusive printed in random order above a 0-to-10 number line. The experimenter asked the child to place a hatch mark to indicate the location of each number on the number line and to write the number they were estimating above the corresponding mark. If the child could not write a number, the experimenter asked what number the child had in mind and wrote the number for the child. The experimenter also indicated that children could estimate the numbers' location in any order they wished and that if at any time they believed that their hatch mark was in the wrong place, they could erase it and place it where they thought it should go. After locating the 10 numbers, children were given an unfilled number line, identical to the initial one, and asked to place their final estimates for each number on that sheet. The children's earlier

estimates remained present while they made their final estimates, which provided a chance for them to think about the estimates and to make any changes that seemed warranted.

The second phase of Session 2 was a posttest in which children in both experimental and control groups estimated the locations of the same numbers on the same type of number line as on the pretest. The time required for Session 2 was similar for the control and experimental groups at a given age, as was the total number of estimates. First and second graders in the experimental group were presented 44 trials versus 48 for peers in the control group; kindergartners in the experimental group were presented 14 trials versus 18 for peers in the control group. Data from the math section of the SAT-9 were obtained 1 to 2 months after the experimental sessions.

Results and Discussion

We first discuss the data from the pretest and then effects of the experimental manipulation. Because all of the analyses parallel those in Experiment 1, we describe them more briefly here.

Accuracy of estimates. A one-way ANOVA on the relation of age or grade to percent absolute error on the pretest replicated the Experiment 1 findings. Kindergartners' estimates were considerably farther from the correct answer than were the estimates of first or second graders (percent absolute error = 24%, 14%, and 10%, respectively), $F(2, 57) = 30.64$, $p < .01$, $\eta^2 = .52$.

Pattern of estimates. As in Experiment 1, group medians were used to test the fit of each function to the estimates produced by children in each grade. Again, the exponential curve fit children's pretest estimates for each number much less well than the linear or logarithmic functions; therefore, it was not included in further analyses of group medians. As shown in Figure 3, the median estimates of the kindergartners were better fit by the logarithmic function than by the linear function ($R^2 = .89$ vs. $.69$), $t(47) = 3.09$, $p < .01$, $d = .45$. In contrast, the median estimates of second graders were better fit by the linear function than by the logarithmic function ($R^2 = .97$ vs. $.85$), $t(47) = 6.33$, $p < .01$, $d = .91$. At the first grade level, the fit of the logarithmic and linear functions did not differ ($R^2 = .94$ and $.92$), $t(47) < 1$.

Analyses of the estimates of individual children yielded a similar pattern of results. The vast majority of children (all but one child) were best fit by the logarithmic or linear models; only these children were examined further. The type of function that fit

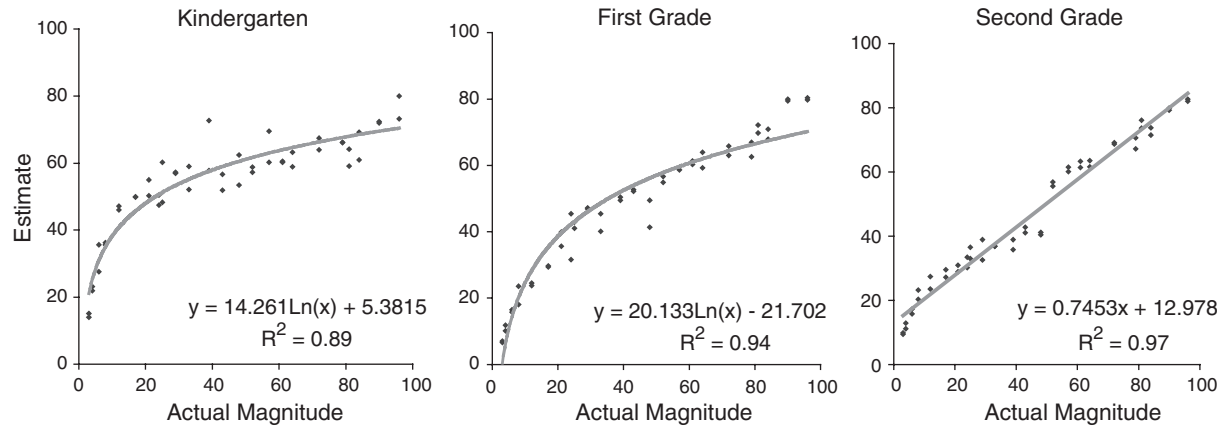


Figure 3. Progression from logarithmic pattern of median estimates among kindergartners (left panel) to linear pattern of estimates among second graders (right panel) in Experiment 2.

the most children varied with age, $\chi^2(2, N = 59) = 17.31, p < .01$. Kindergartners were more likely to be best fit by the logarithmic function (80%) and less likely to be best fit by the linear function (15%) than predicted by chance. In contrast, second graders were more likely to be best fit by the linear function (81%) and less likely to be best fit by the logarithmic function (19%) than predicted by chance. Also as in the analysis of group medians, the estimates of individual first graders were equally likely to be best fit by linear and logarithmic functions (42% and 58% of children, respectively).

An analysis that treated linearity of individual children's estimates as a continuous variable provided converging evidence. A one-way ANOVA indicated that the fit of the linear function to individual children's estimates increased with grade, $F(2, 57) = 37.30, p < .01, \eta^2 = .57$. The linear function provided a better fit to second graders' estimates than to those of first graders (mean $R^2 = .86$ vs. $.68$) and a better fit to first graders' estimates than to those of kindergartners (mean $R^2 = .68$ vs. $.31$).

In addition to estimates becoming more linear with age and experience, the slopes of the best fitting linear function also increased. The mean slope of individual kindergartners' estimates was lower than those of first and second graders, which did not differ (mean slopes = $.38, .62$, and $.72$, respectively), $F(2, 57) = 15.68, p < .01, \eta^2 = .36$.

Variability of estimates. A one-way ANOVA on the relation of age or grade to the mean absolute difference between the two estimates of each number that each child provided indicated that as in Experiment 1, variability decreased with grade, $F(2, 57) = 34.14, p < .01, \eta^2 = .55$. Kindergartners' estimates were more variable than those of first graders, and first

graders' estimates more variable than those of second graders (mean difference = $15.3, 9.1$, and 6.7 , respectively). As in Experiment 1, numerical magnitude was unrelated to the variability of estimates, accounting for 3%, 8%, and 6% of variance among kindergartners, first graders, and second graders, respectively.

Linearity, variability, and accuracy of estimates. Regression analyses of the separate contributions of the linearity and variability of each child's estimates to that child's estimation accuracy, beyond the contribution of the other variable, indicated that as in Experiment 1, degree of linearity was the sole source of the relation. Linearity invariably added at least 32% to the percent absolute error that could be accounted for by variability, but variability never added more than 2% to the variance that could be accounted for by linearity (Table 2). Thus, despite the parallel age trends of variability and percent absolute error, increased linearity entirely accounted for the improvements in estimation accuracy.

Number-line estimation and math achievement. Tests of the correlation between individual children's estimation accuracy and their performance on the mathematics section of the SAT-9 indicated significant relations at two of the three grade levels. Again partialing out age, the correlation for second graders was $pr(17) = -.76, p < .01$, and for first graders, $pr(15) = -.60, p < .01$. The correlation for kindergartners was not significant but was in the same direction, $pr(17) = -.32$. In all cases, the smaller a child's percent absolute error of estimates, the higher was the child's math achievement test score. Relations between linearity and achievement test scores were present at the same grades: second grade, $pr(17) = .81, p < .01$, and first grade, $pr(15) = .54$,

Table 2
Linearity and Variability as Predictors of Percent Absolute Error: Experiment 2

Analysis	Sample	R^2_{Lin}	$R^2_{\text{Lin} + \text{Var}}$	Added R^2	Significance
1. Linearity entered before variability	All children	.90	.90	0	<i>ns</i>
	Kindergartners	.75	.77	.02	<i>ns</i>
	First graders	.87	.87	0	<i>ns</i>
	Second graders	.81	.81	0	<i>ns</i>
2. Variability entered before linearity		R^2_{Var}	$R^2_{\text{Var} + \text{Lin}}$		
	All children	.58	.90	.32	$p < .01$
	Kindergartners	.18	.77	.59	$p < .01$
	First graders	.49	.87	.38	$p < .01$
	Second graders	.20	.81	.61	$p < .01$

$p < .05$. A nonsignificant trend in the same direction was present for kindergartners, $pr(17) = .29$.

Effects of experimental manipulation. Although we expected that the experimental manipulation might be particularly helpful to kindergartners, the experience proved more confusing than helpful to them. Accuracy of kindergartners in the experimental group actually decreased relative to its level before the experimental manipulation (percent absolute error = 12% on the 0-to-10 pretest and 16% on the experimental task), $t(10) = 2.28$, $p < .05$. Accuracy of estimates of kindergartners in the control group also decreased somewhat but not significantly (percent absolute error = 15% and 18%), $t(8) = 1.08$, *ns*. Because the experimental manipulation did not have the expected effect on kindergartners, their performance was not examined further.

On the other hand, the experimental manipulation had the anticipated effects on the estimates of first and second graders. Percent absolute error of first and second graders in the experimental group decreased from the pretest to the experimental session (11% on the pretest, 8% on the posttest), $t(16) = 2.32$, $p < .05$, $d = .56$. In contrast, among peers in the control condition, percent absolute error was unchanged (13% both times). Similarly, the fit of the linear function increased from the pretest to the experimental task (mean $R^2 = .80$ and $.94$), $t(16) = 4.02$, $p < .01$, $d = .84$. Again, no such difference was present among peers in the control group (mean $R^2 = .75$ and $.73$), $t(19) < 1$.

To determine whether the experience of placing all estimates on a single number line improved first and second graders' estimation on the original one-estimate-per-number-line task, we compared their pretest and posttest performance. Within-subject t tests showed that percent absolute error of first and second graders in the experimental group tended to

improve from the pretest to the posttest (11% vs. 9%), $t(16) = 1.72$, $p = .10$, $d = .42$. In contrast, percent absolute error of children in the control group did not change between pretest and posttest; if anything, it became slightly worse (13% and 14%), $t < 1$.

Similarly, the fit of the linear function to the estimates of individual first and second graders in the experimental group tended to increase from pretest to posttest (mean $R^2 = .80$ and $.85$), $t(16) = 1.82$, $p < .10$, $d = .45$. In contrast, the fit of the linear function to the estimates of peers in the control group did not change from pretest to posttest (mean $R^2 = .75$ and $.73$), $t(19) = .63$, *ns*. Thus, first and second graders' estimates of multiple numbers on a single number line were more accurate than their pretest estimates of a single number on each number line, and the experience of estimating the positions of multiple numbers on a single number line tended to improve their later performance on the original estimation task.

General Discussion

Results from Experiment 2, like those from Experiment 1, supported the hypotheses about developmental changes and individual differences that motivated the study. The predicted developmental sequence emerged in both experiments: Almost all kindergartners generated logarithmic patterns of estimates, most second graders generated linear patterns of estimates, and first graders split evenly between the two patterns. The two experiments also told a consistent story regarding the contributions of changes in the linearity and variability of estimates to these changes in accuracy. The timing of improvement in both linearity and variability resembled the timing of improvement in estimation accuracy, but only changes in linearity contributed

independent variance to the changes in accuracy. Findings about individual differences also were consistent across the two experiments. In each case, controlling for age and grade, individual differences in number-line estimation were strongly related to individual differences in math achievement, especially for first and second graders.

Results of Experiment 2 also indicated that, for first and second graders, estimating the locations of multiple numbers on a single number line, and being asked to correct whatever errors were evident, produced more accurate and more linear estimates than the original one-location-per-number-line task. The experience also improved subsequent performance on the original task. These findings have several implications, both for understanding the development of estimation and for understanding cognitive development more generally.

Interpreting Logarithmic-Estimation Patterns

In both experiments, kindergartners' estimates clearly fit a logarithmic function. However, before concluding that this logarithmic data pattern reflected an underlying logarithmic representation of numbers, it seems important to address several questions: Could the logarithmic pattern of estimates have masked an underlying linear representation or perhaps two separate linear representations, one for small numbers and one for large numbers, rather than a logarithmic representation? If the logarithmic data pattern did imply a logarithmic representation, was this representation used on all trials or only on some? Finally, what might the contents of a logarithmic representation of numbers be?

First, consider whether the logarithmic data pattern was evidence for an underlying logarithmic representation of numbers. As Surber (1984) noted, any estimate requires not only assigning values to stimuli but also mapping the values onto the response scale. Thus, kindergartners' logarithmic pattern of estimates might have obscured an underlying linear representation of numbers. This could have occurred in three ways: (a) if kindergartners were confused about the number line or how their representations of numerical magnitudes could be mapped onto it, (b) if kindergartners used a logarithmic mapping function to map a linear representation of numbers onto the number line, or (c) if oversampling of numbers at the low end of the distribution led kindergartners to spread out their estimates at that end.

Although each of these interpretations was possible, several aspects of the data argued against each of them. At least three types of data would be difficult to explain if children did not understand the number-line response format or how their representation could be mapped onto it. One source of evidence was the high absolute fit of the logarithmic function to kindergartners' estimates (75% and 89% of the variance in the two experiments); confusion does not ordinarily produce such systematic patterns of data. Second, the large majority of individual kindergartners (81% and 80% in the two experiments) showed a logarithmic pattern of estimates, again not what would be expected if children were confused. Third, in Siegler and Opfer (2003), on two number-line tasks that were identical except for the number at the right end of the line, many second graders who generated linear estimation patterns on the 0-to-100 lines generated logarithmic patterns on 0-to-1,000 lines. The mapping task was identical in the two cases: Why would the same child be able to map a linear representation onto the number line in one case but not the other?

Several considerations also argued against the possibility that a logarithmic mapping function might have been superimposed onto a linear representation of numbers to yield a logarithmic pattern of estimates. The Siegler and Opfer (2003) finding is again relevant: Why, in a situation in which all mapping requirements were identical, would a child use a linear mapping function to generate responses on a 0-to-100 line but a logarithmic function to generate responses on a 0-to-1,000 line? In addition, previous studies that have directly examined preschoolers' mapping function have concluded that it probably is linear (e.g., Cuneo, 1982). Moreover, a wide range of research with children, adults, and nonhuman animals, using a wide range of response formats and measures, has produced evidence of logarithmic representations of numerical magnitude (Banks & Hill, 1974; Dehaene, 1997).

Oversampling of numbers at the low end of the distribution also seems insufficient to explain the kindergartners' logarithmic pattern of estimates. The oversampling account does not explain why, if both kindergartners and second graders based estimates on linear representations, the kindergartners would have been affected by the skew distribution but the second graders would not have been. The account also does not explain why, in Siegler and Opfer (2003), second graders were able to resist the distribution bias for numerical estimates on the 0-to-100 scale but not the 0-to-1,000 scale. In addition, the

procedures that have yielded evidence for such context effects on numerical estimation (e.g., Birnbaum, 1974) differ from the present procedure in at least two important ways. One crucial difference is that the demonstrations of context effects have involved presenting the entire set of numbers before judgments were made, which made the numerical context clear in advance. Another crucial difference is that responses in those studies required participants to assign numbers to qualitative rating categories (e.g., very large, very small), which are inherently relative to the context, rather than requiring absolute judgments of numerical locations on a number line. Still, it would be worthwhile to examine whether the logarithmic to linear shift would appear with unbiased sampling of numbers. Our prediction is that it would.

Another issue regarding the relation between the logarithmic data pattern and the underlying representation was whether kindergartners' pattern of estimates might have reflected two linear representations, one for small numbers and one for large numbers, rather than a single logarithmic representation. Several of the considerations noted previously are relevant to this issue as well. First, the logarithmic function provides a very good fit to the data; considerations of parsimony argue against assuming two separate functions unless there is compelling evidence for them. Second, logarithmic representations of magnitude have been found in a wide range of studies with non-numerical stimuli, such as studies of adults drawing line lengths (Banks & Hill, 1974); the small-large number distinction could not have produced those findings. Third, Figures 1 and 2 indicate that if there were a break point between two linear functions, it would be around 20; there was no obvious reason why kindergartners would view 15 as a small number but 25 as a large number. Thus, there seemed to be little reason to conclude that kindergartners' estimates reflected two separate linear representations.

A third issue regarding the relation between data and representation concerns whether individual children relied on a given representation on all trials or whether they relied on different representations on different trials. The overlapping waves perspective posits that in many situations children rely on different strategies or representations on different trials. However, this is not always the case; for example, most children use a single approach on balance-scale problems (Siegler & Chen, 1998). The present method did not allow assessment of individual children's representations on each trial, which precluded strong conclusions regarding this

issue. Formulation of alternative methods that allow trial-by-trial assessment of representations is needed to resolve this issue.

A fourth issue concerns the content of the representations that underlie logarithmic and linear patterns of estimates. Following Case and Okamoto (1996), we suspect that children's representations of numerical magnitudes include a strong spatial component, such that larger numbers are translated into larger spatial entities. This may be one reason children as young as 5 years could perform the number-line task in reasonable ways after very brief (about 30 s) instructions. Kindergartners might well require longer and more detailed instructions, and perhaps practice as well, to perform many other numerical-to-non-numerical translations, for example, translating numbers into brightness or loudness. Converging evidence comes from findings that visuospatial deficits often accompany deficiencies in numerical processing (Geary, 1994). One way to test the interpretation would be to determine whether experimental manipulations that interfere with spatial processing hinder numerical magnitude processing as well.

Implications Concerning Acquisition of Linear Representations

How do children come to rely increasingly on linear representations for representing numbers in the 0 to 100 range? Gaining experience with numbers in the 20 to 100 range in the latter part of first grade and throughout second grade seems likely to be a large part of the answer. It was this belief, together with the findings of Siegler and Opfer (2003), that led us to hypothesize that reliance on linear representations would increase substantially in this age range.

In addition to experience in school, informal learning activities, in particular playing board games, also may contribute to increased reliance on linear representations in this age or grade range. Case and Griffin (1990) noted that such games are played far more in middle-income homes than in low-income homes and suggested that this difference contributes to differences among children in numerical understanding in the early elementary school years. Building on this logic, Griffin, Case, and Capodilupo (1995) formulated an instructional intervention that emphasized board games; the intervention led to large and durable improvements in math performance on both investigator-designed measures and standardized achievement tests.

The success of this instructional intervention makes a great deal of sense. Playing board games

provides multiple, redundant cues to the meaning of numbers. The larger the number that appears on the spinner or dice, (a) the greater the distance the child's token traverses, (b) the longer the time it takes to move the token to its destination, (c) the greater the number of discrete moves the child makes if moving the token one square at a time, and (d) the greater the number of words the child states if counting the token's movements while moving it. Thus, board games provide children with strongly correlated spatial, temporal, kinesthetic, and verbal or auditory cues to numerical magnitude. The multiple sources of redundant information constitute an ideal support for constructing a linear representation of numerical magnitude.

The present findings also raise the issue of when children first generate linear representations of single-digit numbers. Results from Huntley-Fenner (2001) suggest such representations are produced by age 5 years. In this study, 5-year-olds were presented a number line with the numbers 1 to 20 along it in a linear array (ascending order, evenly spaced). On each trial, children saw 5, 7, 9, or 11 small squares briefly displayed on a screen and then were asked to point to the number along the number line that matched the number of objects. Children's mean estimates increased linearly with set size, suggesting that they relied on a linear representation. The linear spacing of the numbers along the number line seems likely to have been an important support for the linear pattern of responses because it provided spatial cues to support the numerical cues. Nonetheless, the results suggest that under supportive circumstances, 5-year-olds can use linear representations.

The present results also raise issues regarding numerical magnitude representation at the other end of the developmental period. In particular, if presented unfamiliar ranges of numbers, such as 0 to 10,000,000, would adults also show logarithmic patterns of number line estimates? The present analysis suggests that if computation of percentages were precluded, for example, by short time limits, adults would deviate from their usual linear representations and instead generate logarithmic patterns.

Although experience with particular ranges of numbers seems important for promoting reliance on linear representations, a general understanding of the decimal system also seems essential, especially for representing unfamiliar ranges of numbers. Number lines might provide a useful instructional tool for enhancing such understanding. One possibility is that encouraging children to locate a wide range of numbers on number lines, and providing feedback concerning the numbers' actual locations,

might increase understanding of the decimal system and reliance on linear representations of numbers. Consistent with this hypothesis, first and second graders who participated in Experiment 2 tended to generate more linear patterns of estimates after a single trial in which they were asked to locate 10 numbers between 0 and 100 on a single number line and to think about the relations among the numbers' locations.

Another promising instructional strategy that is already sometimes used in schools is to encourage children to locate benchmarks on number lines (e.g., fourth, half, three fourths) and to use those benchmarks to help locate other numerical magnitudes. The fact that adults and older children appear to use such benchmarks in number-line estimation (Siegler & Opfer, 2003) is one argument in favor of this approach. In addition, number lines have several advantages for classroom use, both practical and conceptual. They are easy to generate, can be used to examine understanding of the magnitudes of any range of numbers, and make possible the understanding that all types of numbers are meaningful entities whose magnitude is defined by the decimal system. Thus, experience with number-line estimation may help children understand the meanings of numbers.

Although the experience of placing many numbers of a single number line increased the accuracy of first and second graders' estimates, the same experience did not help kindergartners estimate more accurately. One possible reason is that kindergartners may have interpreted the request to revise their original answers as indicating that those answers were wrong. Without a clear idea of how the answers should be revised, the kindergartners may simply have become confused. Alternatively, fatigue may have been a problem; kindergartners' decreased accuracy from pretest to posttest in both the experimental and control conditions is consistent with this interpretation. Whatever the reason, it was clear that the number-line experience that helped first and second graders did not help kindergartners.

Implications for Understanding Individual Differences

The correlations in both experiments between number-line estimation and math achievement test scores indicate that number-line estimation is far from an isolated task. Instead, consistent with the view of Case and Okamoto (1996), construction of a linear representation of numbers seems crucial to mathematical development in the early elementary school period. Case and his colleagues did not study

number-line estimation, but the social class differences and instructional effects that they reported provide persuasive evidence for the centrality of a linear representation of numerical magnitudes to mathematical development in this period.

The results of both Case and Okamoto (1996) and the present experiments suggest a prediction regarding individual differences: Reliance on a linear representation within a given number range should be related to ability to learn answers to unfamiliar arithmetic problems in that range. The effect should be concentrated on problems with answers in the compressed part of the logarithmic distribution and should be evident in the quality of errors as well as in the rate of learning correct answers. In particular, for problems whose answers are in the compressed part of the logarithmic representation, first graders whose number-line estimation indicates a linear representation should produce errors that are closer to the correct answer than those of peers whose number-line estimates better fit a logarithmic function.

Implications for a General Understanding of Development

The finding that development of numerical representations in the 0 to 100 range between kindergarten and second grade parallels development in the 0 to 1,000 range between second and sixth grades is reminiscent of a general theme in the writings of classical developmental theorists, in particular, Piaget (1954), Vygotsky (1934/1962), and Werner (1957). All three theorists hypothesized that development is marked by parallel patterns of changes at different ages and over different time spans. Although this proposal has not received much attention in recent years, data from contemporary research provides reason to suspect that the classic theorists were right.

One type of parallel that has emerged in contemporary research involves age-related changes at different levels of complexity in the same domain. The present findings and those of Siegler and Opfer (2003) regarding estimation provide one example. Findings from arithmetic provide another. Learning of single-digit multiplication, which in the United States starts around age 8, shows many similarities to learning of single-digit addition, which starts years earlier. In both tasks, individual children tend to use between three and five strategies, choices among the strategies are adaptive from early in the acquisition process, development involves increasing reliance on more advanced procedural strategies and on retrieval from memory, and adults continue to use procedural strategies on around 20% of trials rather than always relying on retrieval (Ashcraft, 1992;

Geary, 1994; Geary & Wiley, 1991; LeFevre, Bisanz, et al., 1996; LeFevre, Sadesky, & Bisanz, 1996; Siegler, 1987, 1988; Siegler & Shrager, 1984). Such parallels are in no way limited to numerical tasks or even to cognitive development. Karmiloff-Smith (1992) has documented numerous parallels in her representational redescription model, in areas including development of grammar, storytelling, and map-making skills. Similarly, Adolph (1997) found that several aspects of strategy choices in motor activity that emerge in the context of crawling down ramps later repeat themselves when children become able to walk down the ramps.

Another common parallel occurs between early patterns of accuracy and later patterns of solution times. Again, findings from arithmetic are illustrative. Frequency of errors of children just learning to multiply have been found to be more predictive of adult solution times than any structural predictor, such as the product, sum, or square of the sum (Siegler, 1988). Similar findings have emerged in addition (Campbell & Graham, 1985; Groen & Parkman, 1972; Siegler & Shrager, 1984). Again, the parallels are not limited to arithmetic. For example, in language development, age of acquisition of a word is highly predictive of adults' naming time for the word; the effect is present beyond cumulative frequency, prototypicality, phoneme length, number of letters, and imaginability (Moore, Valentine, & Turner, 1999; Morrison, Ellis, & Quinlan, 1992).

A third parallel involves similarities in acquisitions over different time grains. Strong parallels exist between the changes over years observed in cross-sectional and longitudinal studies and the changes over days or weeks observed in microgenetic studies in which children have a denser exposure to relevant experience. As in the previous parallels, relevant data come from studies of arithmetic. Cross-sectional and microgenetic studies of single-digit addition show the same strategies being discovered in the same order, the same slow generalization of newly discovered strategies, and the same shift toward increasing use of retrieval (Siegler, 1987; Siegler & Jenkins, 1989). Similar acquisition patterns also emerge when adults learn to solve letter arithmetic problems, such as $b + e = g$ (Logan & Klapp, 1991; Zbrodoff & Logan, 1986). Recent reviews of microgenetic studies from diverse areas of cognitive development also have noted extensive parallels with findings from traditional cross-sectional studies in the strategies that emerge, the order in which they emerge, the pervasiveness of strategic variability, and the persistence of older inefficient strategies after discovery of more efficient approaches (Kuhn,

1995; Miller & Coyle, 1999; Siegler, 2000). Thus, the message of the classic theorists that development is marked by parallel changes at different ages and over different time spans may be a useful heuristic for modern students of cognitive development as well.

References

- Adolph, K. E. (1997). Learning in the development of infant locomotion. *Monographs of the Society for Research in Child Development*, 62(3, Serial No. 251).
- Assessment Resource Center. (2002). SAT-9 retrieved May 15, 2002, from <http://arc.missouri.edu/k12/sat9.html#dott>.
- Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition*, 44, 75–106.
- Ashcraft, M. H., & Christy, K. S. (1995). The frequency of arithmetic facts in elementary texts: Addition and multiplication in Grades 1–6. *Journal for Research in Mathematical Education*, 26, 396–421.
- Banks, W. P., & Hill, D. K. (1974). The apparent magnitude of number scaled by random production. *Journal of Experimental Psychology Monograph*, 102, 353–376.
- Birnbaum, M. H. (1974). Using contextual effects to derive psychophysical scales. *Perception & Psychophysics*, 15, 89–96.
- Brannon, E. M., Wusthoff, C. J., Gallistel, C. R., & Gibbon, J. (2001). Numerical subtraction in the pigeon: Evidence for a linear subjective number scale. *Psychological Science*, 12, 238–243.
- Campbell, J. I., & Graham, D. J. (1985). Mental multiplication skill: Structure, process, and acquisition. *Canadian Journal of Psychology Special Issue: Skill*, 39, 338–366.
- Case, R., & Griffin, S. (1990). Child cognitive development: The role of central conceptual structures in the development of scientific and social thought. In C. A. Hauert (Ed.), *Developmental psychology: Cognitive, perceptuo-motor, and neuropsychological perspectives* (pp. 193–230). Amsterdam: Elsevier Science.
- Case, R., & Okamoto, Y. (1996). The role of conceptual structures in the development of children's thought. *Monographs of the Society for Research in Child Development*, 61, Nos. 1–2 (Serial No. 246).
- Case, R., & Sowder, J. T. (1990). The development of computational estimation: A neo-Piagetian analysis. *Cognition and Instruction*, 7, 79–104.
- Cohen, R., Weatherford, D. L., Lomenick, T., & Koeller, K. (1979). Development of spatial representations: Role of task demands and familiarity with the environment. *Child Development*, 50, 1257–1260.
- Cuneo, D. O. (1982). Children's judgments of numerical quantity: A new view of early quantification. *Cognitive Psychology*, 14, 13–44.
- Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. New York: Oxford University Press.
- Dowker, A. (in press). *Individual differences in arithmetic: Implications for psychology, neuroscience, and education*. Psychology Press.
- Geary, D. C. (1994). *Children's mathematical development: Research and practical implications*. Washington, DC: American Psychological Association.
- Geary, D. C., & Wiley, J. G. (1991). Cognitive addition: Strategy choice and speed-of-processing differences in young and elderly adults. *Psychology and Aging*, 6, 474–483.
- Gibbon, J., & Church, R. M. (1981). Time left: Linear versus logarithmic subjective time. *Journal of the Experimental Analysis of Behavior*, 7, 87–107.
- Griffin, S. A., Case, R., & Capodilupo, S. (1995). Teaching for understanding: The importance of central conceptual structures in the elementary school mathematics curriculum. In A. McKeough, J. Lupart, & A. Marini (Eds.), *Teaching for transfer: Fostering generalization in learning* (pp. 123–152). Hillsdale, NJ: Erlbaum.
- Groen, G. J., & Parkman, J. M. (1972). A chronometric analysis of simple addition. *Psychological Review*, 79, 329–343.
- Hamman, M. S., & Ashcraft, M. H. (1986). Textbook presentations of the basic addition facts. *Cognition and Instruction*, 3, 173–192.
- Hecox, K. E., & Hagen, J. W. (1971). Estimates and estimate-based inferences in young children. *Journal of Experimental Child Psychology*, 11, 106–123.
- Hiebert, J., & Wearne, D. (1986). Procedures over concepts: The acquisition of decimal number knowledge. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 199–223). Hillsdale, NJ: Erlbaum.
- Huntley-Fenner, G. (2001). Children's understanding of number is similar to adults' and rats': Numerical estimation by 5-7 year-olds. *Cognition*, 78, B27–B40.
- Joram, E., Subrahmanyam, K., & Gelman, R. (1998). Measurement estimation: Learning to map the route from number to quantity and back. *Review of Educational Research*, 68, 413–449.
- Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.
- Kuhn, D. (1995). Microgenetic study of change: What has it told us? *Psychological Science*, 6, 133–139.
- LeFevre, J. A., Bisanz, J., Daley, K. E., Buffone, L., Greenham, S. L., & Sadesky, G. S. (1996). Multiple routes to solution of single-digit multiplication problems. *Journal of Experimental Psychology: General*, 125, 284–306.
- LeFevre, J. A., Greenham, S. L., & Naheed, N. (1993). The development of procedural and conceptual knowledge in computational estimation. *Cognition and Instruction*, 11, 95–132.
- LeFevre, J. A., Sadesky, G. S., & Bisanz, J. (1996). Selection of procedures in mental addition: Reassessing the problem-size effect in adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 216–230.
- Logan, G. D., & Klapp, S. T. (1991). Automatizing alphabet arithmetic: 1. Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 17, 179–195.

- Miller, P. H., & Coyle, T. R. (1999). Developmental change: Lessons from microgenesis. In E. K. Scholnick, K. Nelson, S. A. Gelman, & P. H. Miller (Eds.), *Conceptual development: Piaget's legacy* (pp. 209–239). Mahwah, NJ: Erlbaum.
- Moore, V., Valentine, T., & Turner, J. (1999). Age-of-acquisition and cumulative frequency have independent effects. *Cognition*, 72, 305–309.
- Morrison, C. M., Ellis, A. W., & Quinlan, P. T. (1992). Age of acquisition, not word frequency, affects object naming, not object recognition. *Memory & Cognition*, 20, 705–714.
- Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books.
- Piaget, J., & Inhelder, B. (1956). *The child's conception of space*. London: Routledge and Kegan Paul.
- Schlottmann, A., & Anderson, N. H. (1994). Children's judgments of expected value. *Psychological Psychology*, 30, 56–66.
- Sekuler, R., & Mierkiewicz, D. (1977). Children's judgments of numerical inequality. *Child Development*, 48, 630–633.
- Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General*, 116, 250–264.
- Siegler, R. S. (1988). Strategy choice procedures and the development of multiplication skill. *Journal of Experimental Psychology: General*, 117, 258–275.
- Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.
- Siegler, R. S. (2000). The rebirth of children's learning. *Child Development*, 71, 26–35.
- Siegler, R. S., & Chen, Z. (1998). Developmental differences in rule learning: A microgenetic analysis. *Cognitive Psychology*, 36, 273–310.
- Siegler, R. S., & Jenkins, E. A. (1989). *How children discover new strategies*. Hillsdale, NJ: Erlbaum.
- Siegler, R. S., & Opfer, J. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, 14, 237–243.
- Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In Sophian, C. (Ed.), *The origins of cognitive skills* (pp. 229–293). Hillsdale, NJ: Erlbaum.
- Sowder, J. T. (1992). Estimation and number sense. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 371–389). New York: Macmillan.
- Sowder, J., & Wheeler, M. (1989). The development of concepts and strategies used in computational estimation. *Journal for Research in Mathematics Education*, 20, 130–146.
- Surber, C. F. (1984). Issues in using quantitative rating scales in developmental research. *Psychological Bulletin*, 95, 226–246.
- Vygotsky (1962). *Thought and language* (E. Hanfmann & G. Vakar, Eds. and Trans.). Cambridge, MA: MIT Press. (Original work published 1934).
- Werner, H. (1957). The concept of development from a comparative and organismic point of view. In D. B. Harris (Ed.), *The concept of development: An issue in the study of human behavior* (pp. 125–148). Minneapolis: University of Minnesota Press.
- Zbrodoff, N., & Logan, J. (1986). On the autonomy of mental processes: A case study of arithmetic. *Journal of Experimental Psychology: General*, 115, 118–130.