Microgenetic Changes in Representations of Numerical Magnitude

John E. Opfer (opfer.7@osu.edu)
The Ohio State University, 214 Townshend Hall, 1885 Neil Avenue
Columbus, OH 43210

Robert S. Siegler (rs7k@andrew.cmu.edu)
Carnegie Mellon University, 331 Baker Hall, 5000 Forbes Ave.
Pittsburgh, PA 15213

Abstract

We examined short-term changes in children's numerical estimation, the representational changes that gave rise to improved estimates, and the hypothesis that experiences that highlight discrepancies between logarithmic and linear representations of numerical magnitude result in the greatest changes in estimation accuracy.

Keywords: Number representation; microgenetic studies; cognitive development

Introduction

Numerical estimation is a pervasive process, both in school and in everyday life, but also a process that children find difficult. Whether estimating distance (Cohen et al., 1979), number of objects (Hecox & Hagen, 1971), answers to arithmetic problems (LeFevre, Greenham, & Naheed, 1993), or locations of numbers on number lines (Siegler & Opfer, 2003), 5- to 10-year-olds' estimation is highly inaccurate. The poor quality of children's performance, and the positive relation between estimation proficiency and math achievement (Dowker, 2003; Siegler & Booth, 2004), have led educators to assign a high priority to improving estimation for many years (e.g., NCTM, 1980, 2001). Despite this prolonged effort, most children's estimation skills continue to be poor (Siegler & Booth, 2005).

One likely reason for failures to help children estimate more accurately is that little is known about the representations and processes that underlie estimation, much less about how they develop or what types of experiences produce improvement in them. To help address these limitations of current understanding, we examine in this study how children's choice of representations contributes to their difficulties with estimation. We also propose and test the log discrepancy hypothesis, which predicts the relative effectiveness of different types of experiences for improving children's estimates.

Our account of the development of estimation begins with the assumption that 5- to 10-year-olds' difficulties are due in large measure to inappropriate choices of numerical representation. Specifically, although children in this age range appear to possess multiple representations of numerical magnitudes, they often use an early-developing logarithmic representation (a representation within which the magnitudes denoted by numbers increase logarithmically) in situations where accurate estimation requires use of a linear representation (a representation within which the magnitudes associated with numbers increase linearly).

This analysis raises the issue of how children come to change their representations of numerical magnitude and to use linear representations in situations in which they once used logarithmic ones. It seems likely that over the course of development, children encounter information that does not match their logarithmic representation of numerical magnitudes. If children already know and use linear representations in some numerical contexts, such experiences may lead them to extend the linear representation to numerical ranges where they previously used logarithmic representations.

This logic suggests the *log discrepancy hypothesis*: Experiences should promote extensions of linear representations to new numerical contexts to the extent that the experiences highlight discrepancies between logarithmic and linear representations of numerical magnitudes and clarify the appropriateness of the linear representation. The discrepancy between logarithmic and linear representations of values on a 0-1,000 number line (with both functions constrained to pass through 0 and 1000) is illustrated in Figure 1.
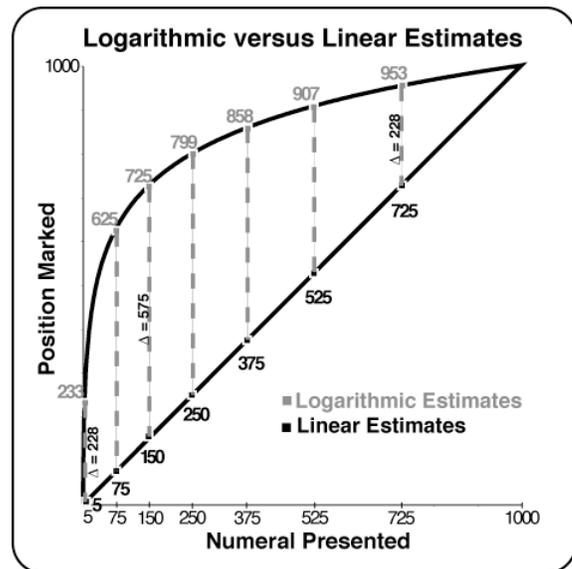

Figure 1: Discrepancy between a logarithmic and linear representation of numeric values on a 0-1,000 number line is greatest at 150; the discrepancies for 5 and 725 are equal to each other and about half as great as that at 150.

If this hypothesis is correct, and improvements in estimation stem from a switch in representations, then changes in patterns of estimates may occur abruptly rather than gradually, and across a broad range of numerical values rather than being local to the numerical range on which feedback is given.

The present study was designed to test the log discrepancy hypothesis regarding how children acquire more advanced representations of numerical magnitudes. The experimental design did this by examining changes in children's number line estimates in response to feedback on numbers around 150, 5, or 725, or in response to answering the same problems without feedback. These experimental conditions allowed tests of the predictions that feedback on numbers around 150 would elicit the largest and quickest change (because it is the area of maximum discrepancy between logarithmic and linear representations), that the change would involve a broad range of numbers and would occur abruptly rather than gradually (because the change involved a choice of a different representation, rather than a local repair to the original representation), and that regardless of the feedback condition, the greatest change would occur on the numbers where logarithmic and linear functions differed by the greatest amount, rather than on the numbers around which children received feedback (again because change was hypothesized to involve substituting linear representations for logarithmic ones).

The design of the current study also allowed us to learn about five key dimensions of cognitive change: the source, rate, path, breadth, and variability of change. These dimensions have proved useful in describing cognitive change in a number of previous contexts (Siegler, 1996). To test the effects of different *sources* of change, we compared the amount of improvement elicited by the four experimental groups. To examine the *rate* of change, we measured how many feedback problems children required in each condition before they adopted a linear representation. To learn about the *path* of change, we tested whether children showed an abrupt shift from a logarithmic pattern to a linear pattern of estimates or whether they progressed from a clear logarithmic pattern to a pattern intermediate between the two functions to a clear linear pattern. To investigate the *breadth* of change, we tested whether amount of change in children's estimates for particular numbers was best predicted by proximity of those numbers to the feedback set or whether the greatest change occurred on the numbers where the discrepancy between the logarithmic and linear representations was greatest, regardless of the feedback set that children received. Finally, to enhance understanding of the *variability* of change, we tested the prediction that children whose pretest estimates adhered most closely to a logarithmic function would learn more than other children and would adhere more closely to the linear function on the posttest.

Method

Participants

Participants included 61 second graders (mean age = 8.2, SD = 0.6) whose estimates in a screening task indicated a better

fitting logarithmic than linear function. A female research assistant served as experimenter.

Task

Each problem consisted of a 25 cm line, with the left end labeled "0," the right end labeled "1,000,"and the number to be estimated appearing 2 cm above the center of the line. The numbers presented were 2, 5, 11, 18, 27, 34, 42, 56, 67, 78, 89, 100, 111, 122, 133, 147, 150, 156, 163, 172, 179, 187, 246, 306, 366, 426, 486, 546, 606, 666, 722, 725, 731, 738, 747, 754, 762, 818, 878, and 938. These numbers included 7 that were between 3 below and 37 above the focal number for each feedback condition (5, 150, and 725); the purpose was to include enough values for the seven feedback problems that children received in each feedback condition.

Design and Procedure

Children were randomly assigned to four experimental conditions: 150-feedback, 5-feedback, 725-feedback, or no-feedback. Children in all four groups completed the number-line estimation task for three trial blocks and a posttest. For children in the three feedback groups, each trial block included a feedback phase and a test phase. The feedback phase included one or three items on which children received feedback; the test phase included 10 items on which children did not receive feedback. Children in the no-feedback group received the same number of estimation trials but always with no feedback. On the posttest, children in all four groups were presented the same 22 problems without feedback as in screening. The children's estimates in screening provided pretest data, which was used as a point of comparison for their subsequent performance.

The only way in which the treatment of children in the three feedback groups differed was in the numbers whose positions they were asked to estimate during the feedback phases. Participants in the 150-feedback group were asked to mark the position of 150 on the first trial block, to mark the positions of 3 numbers from 147–187 on the second trial block, and to mark the positions of a different 3 numbers from 147–187 on the third trial block. Participants in the 5-feedback group were asked to mark the position of 5 on the first trial block, of 3 numbers from 2 – 42 on the second trial block, and of a different 3 numbers from 2 – 42 on the third trial block. Participants in the 725-feedback group were asked to mark the position of 725 on the first trial block, of 3 numbers from 722–762 on the second trial block, and of a different 3 numbers from 722-762 on the third trial block. One-third of the children in the no-feedback group were presented the same problems as the children in the 150-feedback group, one-third were presented the same problems as the children in the 5-feedback group, and one-third were presented the same problems as the children in the 725-feedback group.

The feedback procedure was as follows. On the first feedback problem, children were told, "After you mark where you think the number goes, I'll show you where it really goes, so you can see how close you were." After the child answered, the experimenter took the page from the child and superimposed on the number line a 25 cm ruler (hidden from

the child) that indicated the location of every 10th number from 0-1,000. Then the experimenter wrote the number corresponding to the child's mark ($N_{estimate}$) above the mark, and indicated the correct location of the number that had been presented (N) with a hatch mark. After this, the experimenter showed the corrected number line to the child. Pointing to the child's mark, she said, "You told me that N would go here. Actually, this is where N goes (pointing). The line that you marked is where $N_{estimate}$ actually goes." When children's answers deviated from the correct answer by no more than 10%, the experimenter said, "You can see these two lines are really quite close. How did you know to put it there?" When children's answers deviated from the correct answer by more than 10%, the experimenter said, "That's quite a bit too high/too low. You can see these two lines [the child's and experimenter's hatch marks] are really quite far from each other. Why do you think that this is too high/low for N?"

Results and Discussion

Source of change

We first examined the source of change, the experiences that set the change in motion. To determine whether the particular experience that children received during the feedback phase influenced the degree to which their estimates came to follow a linear function, we compared pretest and posttest performance for the four experimental conditions. In particular, we performed regression analyses on the fit between the children's median estimates for each number and the best fitting logarithmic and linear functions on the pretest and on the posttest.

As shown in Figure 2, on the pretest, all four groups of second graders' median estimates for each number were better fit by the logarithmic function than by the linear one. The precision of the fit of the logarithmic function, and the degree of superiority to the linear function, was similar across the four conditions (5-feedback: log R^2 = .95, lin R^2 = .71, t [21] = 2.71, p < .05; 150-feedback: log R^2 = .95, lin R^2 = .72, t [21] = 2.46, p < .05; 725-feedback: log R^2 = .93, lin R^2 = .68, t [21] = 2.54, p < .05; no-feedback: log R^2 = .94, lin R^2 = .64, t [21] = 2.39, p < .05).

In contrast, the four groups differed considerably in their posttest estimation patterns. Children in the no-feedback group continued to generate estimates that fit the logarithmic function better than the linear one (log R^2 = .90, lin R^2 = .61, t [21] = 2.78, p < .05). Children in the 5- and 725-feedback groups generated posttest estimates for which the fit of the linear function was somewhat, but not significantly, greater than that of the logarithmic function (5-feedback: lin R^2 = .92, log R^2 = .80, t [21] = 1.87, ns; 725-feedback: lin R^2 = .91, log R^2 = .84, t [21] = 1.18, ns). Finally, children in the 150-feedback group generated estimates that fit the linear function significantly and substantially better than the logarithmic one (lin R^2 = .95, log R^2 = .74, t [21] = 2.40, p < .05). This pattern of changes was consistent with the prediction of the log discrepancy hypothesis; the largest change came in response to feedback on problems where the logarithmic and linear functions were most discrepant.
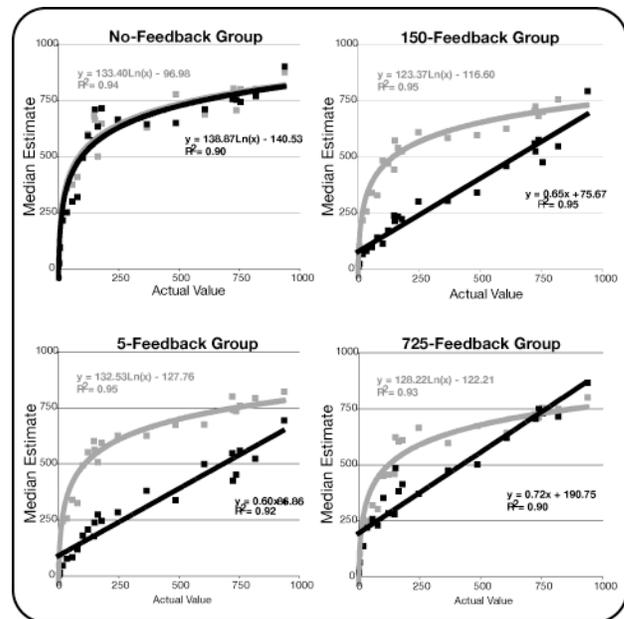

Figure 2: Best fitting functions for pretest (light colored) and posttest (dark colored) median estimates.

A comparison of pretest-posttest improvements in percent absolute error also revealed differences in improvements among the four conditions, F(3, 87) = 2.82, p < .05. On this measure, performance of children in all three feedback conditions improved more than performance of children in the no-feedback control, with no significant differences among the feedback groups.

Rate of change

To examine the rate of change under the four experimental conditions, we compared pretest estimates to estimates given during the no-feedback portion of each trial block during training. We assigned a 1 to the trial blocks of each child that were best fit by the linear function and a 0 to the trial blocks that were best fit by the logarithmic function. The key prediction was that training group and trial block would interact, with the interaction due to children learning fastest in the 150-feedback group and slowest in the no-feedback group.

A 4 (training group: 5-feedback, 150-feedback, 725-feedback, or no-feedback) X 4 (trial block: pretest, 1, 2, 3) repeated-measures ANOVA indicated effects for training group, F (3, 57) = 13.50, p < .001, for trial block, F (3, 171) = 26.36, p < .001, and for the interaction between the two variables, F (9, 171) = 3.83, p < .001. The linear function more frequently fit the estimates of children in the 150-feedback group (60% of trial blocks) than the estimates of children in the no-feedback group (3% of trial blocks, p < .001), 5-feedback group (29% of trial blocks, p < .001) or 725-feedback group (37% of trial blocks, p < .05). The linear function was also the better fitting equation more often for the 5- and 725-feedback groups than for the no-feedback group (p's < .01). The effect of trial block was due to the linear function providing the better fit

more often on trial blocks 1, 2, and 3 (38%, 39%, and 48%, respectively) than on the pretest (0%, p's < .001).

The interaction between training group and trial block (Figure 3) reflected different rates of learning in the four groups. On the pretest, there were no differences among groups in the percentage of children for whom the linear function provided the better fit (it was 0% in all cases). On trial block 1, the linear function fit more children's estimates in the 150-feedback group than in the no-feedback group (p < .001), 5-feedback group (p < .005), or 725-feedback group (all p's < .05). What this meant was that the superiority of the 150-condition for promoting learning manifested itself after feedback on a single estimate. Feedback on this one trial increased the percentage of children for whom the linear function provided the better fit from 0% on the pretest to 85% on the test phase of trial block 1. The linear model also fit more children's estimates in the test phase of trial block 1 among children in the 5- and 725-feedback groups than among children in the no-feedback group (33% and 40% versus 0%, p's < .05).
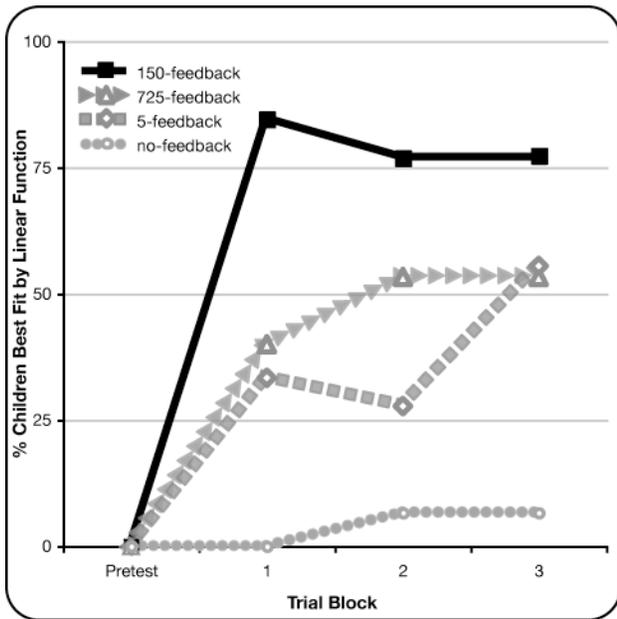

Figure 3: Trial block-to-trial block changes in percentage of children in each condition whose estimates were best fit by the linear function.

On trial block 2, children in the 150-feedback group continued to generate linear patterns of estimates more frequently than children in the 5-feedback and no-feedback groups (77% versus 28% and 7%, p's < .01). The percentage of children in the 725-feedback group who generated linear estimation patterns also was higher than the percentage who did in the no-feedback group (53% versus 7%, p < .01).

By trial block 3, the differences diminished among the three groups that received feedback. The percentage of children whose estimates were better fit by the linear function did not differ among the 150-feedback group (77%), 725-feedback

group (53%), and 5-feedback group (56%), though all three percentages were higher than that in the no-feedback group (7%, all p's < .005).

Another way of testing whether children in the four experimental conditions differed in how quickly they adopted the linear representation was to compare the number of trial blocks before the linear function first provided the better fit to each child's estimates. For this analysis, we excluded children whose estimates were never better fit by the linear function and children in the no-feedback condition, where only one child ever met that criterion on even a single trial block. The fastest learners, children whose estimates were better fit by the linear model on trial block 1, were assigned a score of 1; the slowest learners, children whose estimates were better fit by the linear model for the first time on the posttest, were assigned a score of 4. An ANOVA indicated a trend toward differences among the three feedback groups in the rate of learning, $F(2, 37) = 2.51$, p < .10. The first trial block on which the linear function provided a better fit occurred earlier in the 150-feedback group than in the 5-feedback group (($M = 1.23$ trial blocks versus 2.07, $t(25) = 2.37$, p < .05). The first trial block on which the estimates of children in the 725-feedback group were better fit by the linear function ($M = 1.72$) did not differ from that in either of the other two feedback groups.

Once a child's estimates were better fit by the linear function on one trial block, the child's estimates generally continued to be better fit by it on subsequent blocks. This was true in all three feedback conditions: 73% of trial blocks for children in the 5-feedback condition, 82% of blocks for children in the 150-feedback condition, and 91% of blocks for children in the 725-feedback condition. Thus, once children adopted the linear representation, they generally continued to use it, regardless of the feedback problems that led to its adoption.

Breadth of change.

To examine the breadth of change in children's estimates, we first examined the percentage of the 22 items on which children in the four groups showed pretest-posttest improvements. Children in the no-feedback condition generated more accurate posttest estimates on only 36% of items (8 of 22.) In contrast, the estimation accuracy of children in the three feedback groups improved on an average of 70% of items, with similar percentages (64%-77%) in the three groups. Thus, feedback produced improvement on a broader range of items than simply performing the estimation task.

The next goal was to identify the range of numbers on which the greatest improvement in estimation accuracy occurred. In particular, we wanted to examine whether improvements in accuracy followed a standard generalization gradient, in which learning decreases with distance from feedback items, or whether the discrepancy between logarithmic and linear representations for each number was the key determinant of improvement, regardless of the particular feedback problems.

We first tested the generalization gradient hypothesis. To do this, we regressed pretest-posttest change in absolute error for each number against the distance between that number and the focal number for each feedback group (5, 150, or 725). Results

of this analysis presented a puzzling pattern. Results for two of the three feedback conditions were consistent with the generalization gradient hypothesis. Percent variance in pretest-posttest improvement accounted for by distance between the feedback and test items was $R^2 = .72$ in the 5-feedback group ($F(1,21) = 50.52$, $p < .001$), and $R^2 = .82$ in the 150-feedback group ($F(1,21) = 91.79$, $p < .001$). However, the relation in the 725-feedback condition was not only much weaker ($R^2 = .33$, $F(1, 21) = 9.87$, $p < .01$) -- it was actually in the opposite direction of that predicted by the generalization gradient hypothesis. That is, in the 725-feedback group, the improvement following feedback was greater for test items that were *further* from the feedback items.

Fortunately, there was a straightforward explanation for this seemingly odd pattern: Improvement in estimation was not a function of distance from the feedback problems but rather of the discrepancy between the logarithmic and linear representations. In all three feedback conditions, the largest improvements occurred for numbers where the discrepancies between the logarithmic and linear representations were greatest (numbers around 150), regardless of how far those numbers were from the numbers on which children received feedback. This pattern emerged most dramatically in comparisons between amount of pretest-posttest improvement on the exact items on which children received feedback and amount of pretest-posttest improvement on the numbers around 150. In the 5-feedback group, pretest-posttest improvements for the numbers on which children had received feedback and that were also on the pretest and posttest (2, 5, and 18) were quite modest (4%, 9%, and 18% of the possible improvement, respectively). The improvements for the numbers around 150 (147, 150, and 163) were noticeably larger (36%, 31%, and 25% of the possible improvement), despite these numbers being further away from the numbers on which feedback had been given. A similar pattern was evident for the 725-feedback group, where improvements for the three values near 150 (16%, 13%, and 17% of the possible improvement) were among the greatest in the group, whereas accuracy on the numbers on which feedback had been given actually showed small decreases (-4%, -1%, and -4%). No comparable comparison for the 150-feedback condition was possible, because in it, the generalization gradient and log discrepancy hypotheses led to identical predictions that numbers around 150 should show especially large improvements (which they did, 21%, 30%, and 25% of the possible improvement.)

To examine the breadth of change in a way that would include all 22 numbers on the pretest and posttest and would also allow tests for all three feedback conditions, we regressed pretest-to-posttest change in accuracy against the discrepancy between logarithmic and linear representations for the to-be-estimated number. To compute the discrepancies between the linear and logarithmic functions, we used the formula y=x for the linear function and $y = 144.761(\ln(x))$, the same equations used in Siegler and Opfer (2003). These equations were chosen so that both functions would pass through 1 and 1,000.

The discrepancy between the logarithmic and linear functions provided an excellent fit to the improvement in all three feedback conditions, and the effect was in the predicted direction in all conditions: in the 5-feedback group, $R^2 = .76$, $F(1, 21) = 64.37$, $p < .001$; in the 725-feedback group, $R^2 = .62$, $F(1, 21) = 32.67$, $p < .001$; and in the 150-feedback group, $R^2 = .73$, $F(1, 21) = 52.69$, $p < .001$. The findings were not attributable to regression to the mean being greatest at the points where the pretest estimates were most discrepant; the parallel analysis for children in the no-feedback group did not show any relation between log-linear discrepancy and pretest-posttest improvement, $R^2 = .12$, $F(1, 21) = 2.84$, *ns*.

To appreciate just how powerful this relationship was, consider the subset of 9 numbers on which the discrepancy between the logarithmic and linear functions was above 500. These 9 numbers, which ranged from 56 to 246, were the items on which the log discrepancy hypothesis predicted the greatest improvement regardless of experimental condition. In both the 5-feedback condition and the 725-feedback condition, all 9 numbers were among the 11 on which children showed the greatest improvement; in the 150-feedback condition, the 9 numbers were exactly the 9 numbers on which improvement was greatest. Again, this was not attributable to regression to the mean. In the no-feedback condition, only 3 of the 11 numbers on which change was most positive (or least negative) were in this range. Particularly striking, children who received feedback on numbers from 722-762, like the other children, showed the greatest improvement on numbers from 56 to 246.

These results suggested two conclusions regarding the breadth of change. First, the change was more than a local repair to children's estimation procedures; improvements in posttest accuracy were not limited to, or even greatest at, the areas of the number line on which children received feedback. Second, the change seemed to entail substitution of a linear representation for a logarithmic one, as indicated by the improvements in estimates being greatest for numbers where the two representations differed by the greatest amount. Analyses of children's path of change lent additional support to these conclusions, as described in the next section.

Path of change.

Children could have moved from a logarithmic to a linear representation via several paths. To examine which path(s) they actually took, we examined trial-block-to-trial-block changes in individual children's estimates. In particular, we identified the first trial block on which the linear function provided a better fit than did the logarithmic function to a given child's estimates on the 10 no-feedback test items, and we labeled it "trial block 0." The trial block immediately before each child's trial block 0 was that child's trial block –1", the trial block before that was the child's "trial block -2" and so on.

These assessments of the trial block on which children's estimates first fit the linear function made possible a backward-trials analysis that allowed us to test alternative hypotheses about the path of change from a logarithmic to a linear representation. One hypothesis was that the path of change entailed gradual, incremental improvements in the linearity of estimates. According to this hypothesis, the fit of the linear model would have gradually increased, and the fit of logarithmic model would have gradually decreased, from Trial

Block -3 to Trial Block +3. In this scenario, Trial Block 0 — the first trial block in which the linear model provided the better fit — would mark an arbitrary point along a continuum of gradual improvement, rather than a point at which children switched from one representation to another.

A second hypothesis was that the path of change involved initial reliance on a logarithmic representation, followed by a period of disequilibrium or confusion, followed by reliance on a linear representation. According to this Piagetian-inspired hypothesis, the fit of the logarithmic model would have been very good initially (e.g., in Trial Blocks -3 and -2). However, feedback would then have confused the child and led to a poor fit of both linear and logarithmic models immediately before the change (i.e., in Trial Block -1). Then the child would resolve the conflict by adopting the linear representation (on Trial Block 0 and thereafter).

A third hypothesis was that the path of change involved a discontinuous switch from a logarithmic to a linear representation, with no intermediate state. This would have entailed no change in the fit of the linear model from Trial Block -3 to -1, a large change from Trial Block -1 to Trial Block 0, and no further change after Trial Block 0.

This third hypothesis fit the data. From Trial Block -3 to -1, there was no change in the fit to children's estimates of either the linear function or the logarithmic function (F's< 1). There also was no change from Trial Block 0 to Trial Block 3 in the fit to children's data of either the linear or the logarithmic function (F's< 1). However, from Trial Block -1 to Trial Block 0, there was a large increase in the fit of the linear function to individual children's estimates, from an average R^2 = .57 to an average R^2 = .80, $F(1, 75) = 25.67$, $p < .001$. Complementarily, there was a decrease from Trial Block -1 to Trial Block 0 in the fit of the logarithmic function to children's estimates, from an average R^2 = .74 to an average R^2 = .64, F $(1, 75) = 4.95$, p < .05. Thus, rather than Trial Block 0 reflecting an arbitrary point along a continuous path of improvement, or reflecting the end of a period of disequilibrium, it seemed to mark the point at which children switched from a logarithmic representation to a linear one.

Variability of change

The log discrepancy hypothesis suggested that children whose initial representations were consistently *logarithmic* might respond to feedback by adopting representations that were more consistently *linear* than would children whose initial representations were less consistently logarithmic. The reason is that the difference between the children's estimates and the feedback they received would be more dramatic, and thus more likely to motivate a shift to the alternative (linear) representation, among children whose initial estimates were most strongly logarithmic.

To test this hypothesis, we correlated percent variance accounted for by the logarithmic function on the pretest with percent variance accounted for by the linear function on the posttest. As hypothesized, the fit of the logarithmic model to each child's pretest estimates predicted the fit of the linear model to the child's posttest estimates (r= .36, $F(1, 45) = 6.75$, $p < .05$). The better the logarithmic model fit the children's

pretest estimates, the better the linear model fit their posttest estimates. This correlation was chiefly evident among children whose posttest estimates were better fit by the linear function (r = .42, $F[1, 29] = 6.02$, $p < .05$); it was not found among children whose posttest estimates were not better fit by the linear function (r = .08, ns). In contrast, the fit of the linear function to children's pretest estimates did not predict the fit of the linear function to their posttest estimates among either children who adopted the linear model (r = .19, ns) or those who did not (r = .32, ns). Thus, the log discrepancy hypothesis yielded counterintuitive but accurate predictions regarding individual differences in learning, as well as the types of feedback that would trigger the largest changes and the types of numbers on which improvements in estimation would be greatest.

References

Cohen, 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.

Dowker, A. (2003). Young children's estimates for addition: The zone of partial knowledge and understanding. In Baroody, A. J., Dowker, A. (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise* (pp. 243-265). Mahwah, NJ: USum Associates, Publishers.

Hecox, K. E.,& Hagen, J. W. (1971). Estimates and estimate-based inferences in young children. *Journal of Experimental Child Psychology, 11*, 106-123.

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.

National Council of Teachers of Mathematics (1980). *An agenda for action: Recommendations for school mathematics of the 1980s.* Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics: Higher standards for our students . . . Higher standards for ourselves.* Washington, DC: National Council of Teachers of Mathematics.

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking.* New York: Oxford University Press.

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

Siegler, R. S. & Booth, J. L. (2005). Development of numerical estimation: A review. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp 197-212). Boca Ratan, FL: CRC Press.

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