Available online at www.sciencedirect.com
ScienceDirect
Cognitive Psychology 55 (2007) 169–195
Cognitive Psychology
www.elsevier.com/locate/cogpsych

Representational change and children's numerical estimation

John E. Opfer [a,*], Robert S. Siegler [b]

[a] *The Ohio State University, USA*
[b] *Carnegie Mellon University, USA*

Accepted 1 September 2006
Available online 25 April 2007

Abstract

We applied overlapping waves theory and microgenetic methods to examine how children improve their estimation proficiency, and in particular how they shift from reliance on immature to mature representations of numerical magnitude. We also tested the theoretical prediction that feedback on problems on which the discrepancy between two representations is greatest will cause the greatest representational change. Second graders who initially were assessed as relying on an immature representation were presented feedback that varied in degree of discrepancy between the predictions of the mature and immature representations. The most discrepant feedback produced the greatest representational change. The change was strikingly abrupt, often occurring after a single feedback trial, and impressively broad, affecting estimates over the entire range of numbers from 0 to 1000. The findings indicated that cognitive change can occur at the level of an entire representation, rather than always involving a sequence of local repairs.

Keywords: Numerical estimation; Representation; Microgenetic method; Conceptual development

1. Introduction

Numerical estimation is a pervasive process, both in school and in everyday life. It also is a process that most children find difficult. Whether estimating distance (Cohen, Weatherford, Lomenick, & Koeller, 1979), amount of money (Sowder & Wheeler, 1989), number of

* Corresponding author.
E-mail address: opfer.7@osu.edu (J.E. Opfer).

0010-0285/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.cogpsych.2006.09.002

discrete 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 overall mathematics achievement (Dowker, 2003; Siegler & Booth, 2004), have led educators to assign a high priority to improving estimation for at least the past 25 years (e.g., NCTM, 1980, 1989, 2000). Despite this prolonged effort, most children's estimation skills continue to be poor (Dowker, 2003; Siegler & Booth, 2005).

This poor skill level is particularly unfortunate because estimation is not only a pervasive activity but also one that plays a central role in a wide range of mathematical activities. Accuracy of numerical estimation correlates quite highly with standardized achievement test performance (Dowker, 2003; Siegler & Booth, 2004) and also with specific numerical processes, such as arithmetic and magnitude comparison (Booth, 2005; Laski & Siegler, 2005). Consistent patterns of individual and developmental differences are also present across numerical estimation tasks. Children who are skillful at one type of numerical estimation tend to be skillful at others, and different numerical estimation tasks show parallel changes at the same ages (Booth & Siegler, 2006). Thus, numerical estimation seems to be a coherent category, one well worth understanding.

In the present study, we attempt to go beyond previous studies of pure numerical estimation to address a fundamental question about it and about development more generally: how does representational change occur? Our approach to this question is based at a general level on overlapping waves theory and at a more specific level on the hypothesis that encountering information that is clearly at odds with existing, non-optimal representations leads to rapid and broad representational change and thus to improved estimation.

1.1. Representational change as an adaptive choice

The theory on which the present research is based, overlapping waves theory, depicts children at any given age as knowing and using a variety of approaches (i.e., strategies, rules, or representations) that compete with one another for use, with each approach being more or less adaptive depending on the problem and situation (Siegler, 1996). With use of these varied approaches across different types of problems, information accumulates about the relative adaptiveness of each approach on specific types of problems and the more adaptive approaches are chosen increasingly (albeit often unconsciously; Siegler & Shrager, 1984). In situations where people do not gain information about the adaptiveness of a strategy or representation, performance tends not to change, leading to older children and adults sometimes continuing to use approaches that are typical of young children.

Overlapping waves theory differs from alternative approaches to representational development such as stage theories (e.g., Bruner, Olver, & Greenfield, 1966), incremental theories (e.g., Brainerd, 1978), and early competence theories (e.g., Gelman & Gallistel, 1978) in its claim that individuals usually know and use multiple strategies and representations, rather than only a single one, and in its emphasis on choices among the alternative approaches. Like stage and early competence theories, it recognizes that cognitive change sometimes involves rapid substitution of one strategy or representation for another; unlike them, such changes are viewed as atypical, rather than as the norm. Abrupt changes are most likely when there are large differences between the accuracy yielded by the new strategy or representation and that possible with older approaches (Siegler, 2006), a generalization that proved important in the present study.

Overlapping waves theory, together with recent empirical findings summarized in Siegler and Booth (2005), suggests that 5- to 10-year-olds' difficulties with numerical estimation are due in large measure to inappropriate choices of numerical representation. Specifically, children in this age range are hypothesized to possess multiple representations of numerical magnitudes, and to 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). An example of using a logarithmic representation would be estimating the difference between $1 and $100 as being larger than the difference between $901 and $1000. In contrast judging the differences in money to be equal would imply the use of a linear representation of numerical value. Children between 5 and 10 years of age are believed to rely on linear representations with small numbers but to only gradually extend the linear representations to larger numbers and to numbers other than integers.

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 (e.g., hearing 150 referring to a relatively small part of 1000 items). If children already apply linear representations in some numerical contexts (e.g., for small numeric ranges), such experiences may lead them to draw analogies between the two contexts and to extend the linear representation to numerical ranges where they previously used logarithmic representations. In other words, experiences that are at odds with their logarithmic representations of numerical magnitudes may lead them to extend a linear representation that they use in small number contexts to large number ones.

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 make clear the appropriateness of the linear representation in the new contexts. If this hypothesis is correct, and improvements in estimation stem from a substitution of one representation for another, 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. Implications of the discrepancy hypothesis extend well beyond estimation; the hypothesis implies that whenever people rely on a less advanced representation in some contexts and a more advanced one in other, structurally parallel ones, experiences that highlight the advantages of the more advanced approach may trigger rapid substitution of one representation for another, at least in that situation. This more general *representational discrepancy hypothesis*, and evidence favoring it, will be discussed later in the article.

The present study tests the log discrepancy hypothesis and its implications in two experiments. Experiment 1 examines whether development between second and fourth grade is greatest in the numerical region that in theory should show the greatest improvement during this period, the region in which the discrepancy between logarithmic and linear representations is greatest. Experiment 2 examines whether providing second graders feedback on problems that in theory should stimulate the greatest improvement (problems on which the discrepancy between logarithmic and linear functions is greatest) do in fact stimulate greater learning across the full range of numbers than feedback on other problems or simple experience with the estimation task. To make clear the logic and evidence that motivated these experiments, we next review evidence about the relation between estimation and representations of numerical magnitudes, present a general theoretical analysis of the development of numerical estimation, and describe how the analyses led to the present experiments.

1.2. Estimation and representations of numerical magnitudes

Recent findings suggest that an important source of children's difficulty with numerical estimation is inappropriate choice of numerical representation. Specifically, children often choose an early-developing logarithmic representation in situations where accurate estimation requires use of a linear representation (Siegler & Opfer, 2003; Siegler & Booth, 2004). The use of logarithmic representations of quantities is widespread among species and age groups from infants to adults (Banks & Hill, 1974; Feigenson, Dehaene, & Spelke, 2004; Holyoak, 1978; Moyer, 1973; Siegler & Opfer, 2003), and for good reason: in a great many situations, such representations are useful. For example, to a hungry animal, the difference between 2 and 3 pieces of food is far more important than the difference between 87 and 88 pieces; for people, the difference between receiving a gift of $1 and $100 is far more important than the difference between receiving a gift of $1,000,001 and $1,000,100. In the formal numerical system, however, magnitudes increase linearly rather than logarithmically. Thus, children's use of logarithmic representations of numerical magnitudes is understandable, but in school and modern life, it can interfere with accurate estimation. According to overlapping waves theory, the inaccuracies produced by the logarithmic representation, together with the extensive experience that children have with some estimation tasks and numerical ranges, set the stage for an age-related trend toward increasing use of the linear representation on those tasks and numerical ranges.

Developmental shifts from a logarithmic to a linear representation have in fact been found between kindergarten and second grade for estimates of numerical locations on 0–100 number lines (Siegler & Booth, 2004) and between second and sixth grade for estimates of numerical locations on 0–1000 lines (Siegler & Opfer, 2003). Thus, when asked to estimate the locations of numbers on number lines with 0 at one end and 100 at the other and no markings in between, most kindergartners produced estimates consistent with a logarithmic function, most second graders produced estimates consistent with a linear function, and about half of first graders produced estimates that were best fit by one function and half by the other (Siegler & Booth, 2004). Similarly, when asked to estimate the locations of numbers on number lines with 0 at one end and 1000 at the other, the large majority of second graders generated logarithmic distributions of estimates, the large majority of sixth graders produced linear distributions, and about half of fourth graders produced estimates that were best fit by each function (Siegler & Opfer, 2003). By second grade, if not earlier, individual children possess both types of representations; almost half of the second graders in Siegler and Opfer (2003) generated a linearly increasing pattern of estimates on 0–100 number lines and a logarithmically increasing pattern on 0–1000 lines.

Findings regarding both developmental and individual differences on a variety of estimation tasks are consistent with the hypothesis that children's difficulties in numerical estimation stem in large part from inappropriate choices of representation. From second to fourth grade, trends toward more linear and less logarithmic patterns of estimates are evident for measurement and numerosity estimation, as well as for number line estimation. Booth and Siegler (2006) found that from second to fourth grade, the mean percentage of variance in individual children's estimates accounted for by the best fitting linear function increased from 66% to 85% on a number line estimation task, from 59% to 83% on a measurement estimation task, and from 57% to 77% on a numerosity estimation task. Individual differences in use of linear representations on the three tasks were highly correlated for both second and fourth graders. The types of experiences that lead to such improvements in children's estimation, and the process through which the change occurs, remain unknown, however. Improving our understanding of this process of representational change is the central purpose of the present study.

1.3. Development of numerical magnitude representations

The findings of Siegler and Opfer (2003) suggest that second graders can represent numerical magnitudes in the 0–100 range either logarithmically or linearly, but that they only represent numerical magnitudes in the 0–1000 range logarithmically (probably due to less familiarity with numerals that represent large numerosities; Dehaene, 1990). Development beyond second grade seems to involve children learning that the linear representation that is useful for the range 0–100 also is useful with larger numerical ranges.

From this perspective, if children who apply linear representations to some numerical ranges learn that their estimates in other numerical ranges are inaccurate, and if accurate estimates would have been predicted by a straightforward extension of the linear representation to the other numerical range, then children are likely to draw an analogy between the two numerical ranges and to map the linear representation onto the new range. For example, if a second grader is shown that her estimate of the position of 150 on a 0–1,000 number line is too high, and also is shown the correct position of 150 within that range, she may draw the analogy "150 is to the 0–1000 range as 15 is to the 0–100 range." This analogy may lead her to choose a linear representation for the 0–1000 range on subsequent estimation problems. If the analogy is drawn at the level of the entire representation (as opposed to being restricted to numbers near 150), such feedback would lead to more accurate estimates for numbers throughout the 0–1000 range, especially numbers where the log and linear representations differ most dramatically. Such a substitution of representations could occur quite quickly, because the linear representation has already been constructed and used in smaller numerical contexts.

What types of experiences would be most likely to stimulate such an analogy? The log discrepancy hypothesis predicts that if children are using a logarithmic representation, then the magnitude of change in their estimates in response to feedback should be positively related to the discrepancy between the logarithmic and linear functions for the problems on which the children receive feedback. Larger discrepancies between children's estimates and the linear function are more likely to provoke the realizations that the underlying representation is wrong and that a new way of thinking about the task is needed. In contrast, smaller discrepancies between estimated and correct values may be attributed to misapplication of a basically correct approach, which may motivate children to try to be more careful rather than to choose a different representation for the entire class of numbers. Estimation experience that does not make clear the superiority of the linear representation would not be expected to evoke substantial change.

The discrepancy between a logarithmic and a linear representation of the values on a 0–1000 number line (with both functions constrained to pass through 0 and 1000) is illustrated in Fig. 1. As the figure shows, the difference in estimates varies as a function of the number presented. The maximum difference occurs at 150, where the logarithmic representation predicts an estimate of 725 and the linear representation predicts an estimate of 150, resulting in a discrepancy of 575 (57.5% of the line). For purposes of comparison, the absolute numerical discrepancy between the estimates predicted by the linear and logarithmic representations of both 5 and 725 is 228 (22.8% of the line).

Thus, feedback that indicated the correct position of 150 on a 0–1000 number line would direct learners' attention to the area where the discrepancy between children's initial (logarithmic) understanding and the correct (linear) understanding is greatest. If the difference between the prediction of the representation and the correct value is an important variable in the probability of extension of the linear representation to the 0–1000 context, then information about the correct placement of 5 or 725 on the number line should be less effective than information about the correct placement of 150. Moreover, the effects of feedback regarding the correct estimate for 5 and for 725 should be equivalent to each other, and feedback about either should promote greater learning than simply performing the estimation task without feedback.

1.4. Issues examined in Experiment 1

Experiment 1 had three major purposes. One was to test whether the greatest improvement between second and fourth grade occurs for numbers around 150. Siegler and Opfer's (2003) stimulus set did not include any numbers in this area—the closest numbers that they presented were 86 and 230—but the theoretical prediction was that the greatest

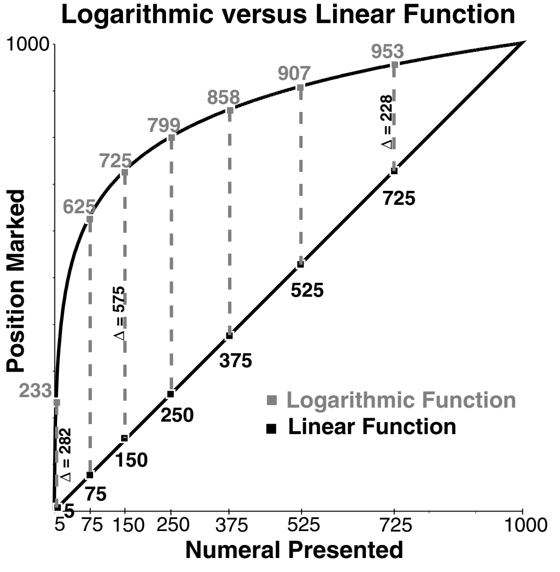

Fig. 1. The discrepancy between a logarithmic and linear representation of numeric values on a 0–1000 number line is greatest at 150; the discrepancies for 5 and 725 are equal to each other and about half as great as that at 150.

improvement with age should come in this area, because it is the area where the logarithmic and linear functions are most discrepant. The second purpose of Experiment 1 was to identify individual second graders whose estimates fit a logarithmic function better than a linear one. This second goal was important because these children subsequently participated in a microgenetic study of the transition from use of a logarithmic representation to use of a linear representation (Experiment 2). The third purpose of Experiment 1 was to replicate and extend Siegler and Opfer's (2003) findings regarding the log to linear shift with a larger range of numbers (the original set included only 12 numbers between 0 and 1000, only 2 of which exceeded 500).

2. Experiment 1: age-related differences in number-line estimation

2.1. Method

2.1.1. Participants

Participants were 93 second graders (mean age = 8.2 years, $SD = 0.6$) and 60 fourth graders (mean age = 10.3 years, $SD = 0.5$). The children attended a suburban school in a middle class area. A female research assistant served as experimenter.

2.1.2. Tasks

Each problem consisted of a 25 cm line, with the left end labeled "0" and the right end labeled "1000." The number to be estimated—2, 5, 18, 34, 56, 78, 100, 122, 147, 150, 163, 179, 246, 366, 486, 606, 722, 725, 738, 754, 818, and 938—appeared 2 cm above the center of the line. These numbers were chosen to maximize the discriminability of logarithmic and linear functions by oversampling the low end of the range, to minimize the influence of specific knowledge (such as that 500 is halfway between 0 and 1000), and to test predictions about the range of numbers where estimates of the two age groups would differ most.

2.1.3. Procedure

Participants were tested in a single session. The items within each scale were randomly ordered, separately for each child, and presented in small workbooks, one problem per page. The experimenter began by saying, "Today we're going to play a game with number lines. What I'm going to ask you to do is to show me where on the number line some numbers are. When you decide where the number goes, I want you to make a line through the number line like this (making a vertical hatch mark)." Before each item, the experimenter said, "This number line goes from 0 at this end to 1000 at this end. If this is 0 and this is 1000, where would you put *N*?"

2.2. Results and discussion

We first compared the fit to second and fourth graders' median estimate for each number that was generated by the best fitting linear and logarithmic functions. As in Siegler and Opfer (2003), the fit of the linear function to children's estimates increased from second to fourth grade, whereas the fit of the logarithmic function decreased. The best fitting logarithmic function fit second graders' median estimates better than did the best fitting linear function (Fig. 2); the logarithmic equation accounted for 95% of variance in median estimates versus 80% for the linear equation, $t(21) = 2.27$, $p < .05$, $d = .65$. In contrast, the

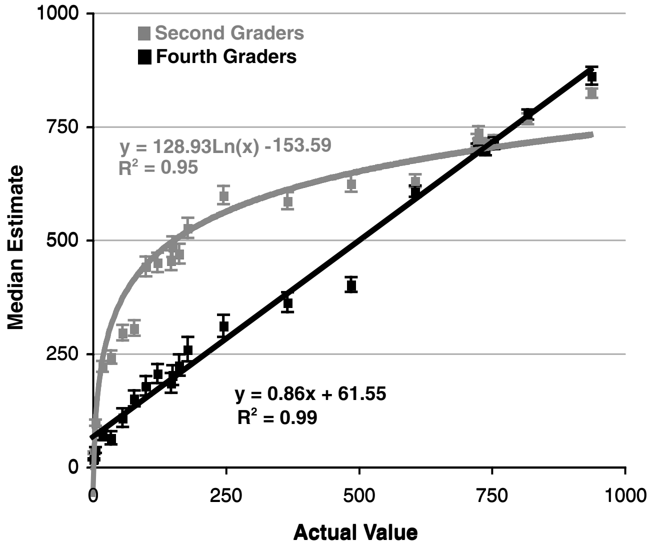

Fig. 2. Experiment 1. Age differences in number-line estimation.

linear function fit fourth graders' median estimates better than did the logarithmic function (99% versus 72% of variance, $t(21) = 5.33$, $p < .01$, $d = 1.60$).

To test whether the group medians reflected individual performance, we regressed individuals' estimates on each task against the predictions of the best fitting linear and logarithmic functions. We assigned a 1 to the model that best fit each participant's estimates and a 0 to the other model.

The function that provided the better fit to individual children's estimates varied with the children's age in ways that mirrored the findings with the group medians, $\chi^2(1) = 30.39$, $p < .001$. The linear function provided the better fit for 34% of second graders and 80% of fourth graders, whereas the logarithmic function provided the better fit for 66% of second graders and 20% of fourth graders.

We also examined whether the greatest age differences in estimates occurred on numbers around 150, where the discrepancy between the logarithmic and linear functions was greatest. This analysis was important because it is possible, for example, for children to have a linear representation with a very high or low slope, thereby affecting where the maximum discrepancy would actually occur. To calculate improvement with age, we first converted the median magnitude estimate for each number (the student's hatch mark) to a numeric value (the linear distance from the "0" mark to the student's hatch mark), then divided the result by the total length of the line, then multiplied the result by 1000, and then calculated the absolute differences between each age group's median for each number and the correct value for the number.

After these calculations, we correlated the absolute numerical distance of each to-be-estimated number from 150 with the decrease in absolute error of estimates between second and fourth grade on that number. Improvement in estimation accuracy proved to be highly correlated with distance from 150: $r(21) = -.80$, $p < .001$; the closer the number to 150, the greater the improvement with age. We then examined estimates for a fixed numerical range (32) around three anchors of interest-150 (where the discrepancy in estimates is greatest between the logarithmic and linear functions), 725 (where the discrepancy is 40% of the discrepancy at 150), and 5 (where the discrepancy is also 40% of the discrepancy at 150). The stimulus set included four numbers in each of these three numerical ranges. As anticipated, improvements in estimation accuracy between second and fourth graders was greater for the four numbers around 150 than for either the four numbers around 725 (mean improvement of 266 versus 18; $t[3] = 37.76$, $p < .01$) or for the four numbers around 5 (mean improvement of 266 versus 101, $t[3] = 3.84$, $p < .05$). Second and fourth graders' estimates around 5 and 725 did not differ significantly from each other ($t[3] = 1.90$, n.s.).

Percent absolute error also tended to improve between second and fourth grade, decreasing from 18% for second graders to 13% for fourth graders, $F(1, 43) = 3.52$, $p < .07$ $d = .56$. This trend reflected the presence of two sub-groups of second graders. The estimates of 34% of second graders best fit the linear representation; these children's percent absolute error did not differ from that of the fourth graders. In contrast, the estimates of the other 66% of second graders best fit the logarithmic function; their percent absolute error was considerably greater than the fourth graders' (23% versus 13%, $F(1, 43) = 7.46$, $p < .01$, $d = .82$).

Thus, Experiment 1 yielded three main findings. First, children's estimates became more linear and less logarithmic between second and fourth grade. These results replicate and extend Siegler and Opfer's (2003) findings regarding age-related changes, with a wider range and greater number of estimated values. Second, the specific numbers that elicited the largest age differences were those predicted by the log discrepancy hypothesis. If children's improvements arose merely from their eliminating random errors in their placement of hatch marks, there would be no reason to expect that age differences should be greatest for numbers around 150; however, the finding was directly predicted by the log discrepancy hypothesis. Third, we were able to identify a large number of second graders whose estimates were best fit by a logarithmic function, which allowed us to examine their acquisition of linear representations in Experiment 2.

3. Experiment 2: the process of representational change

Experiment 2 was designed to test the log discrepancy hypothesis regarding how children acquire more advanced representations of numerical magnitudes. In particular, we examined 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 this 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 reasoning underlying these predictions, especially the second one, merits some discussion. All three predictions stem from the theoretical view that children's estimates of each number reflect a coherent representation of the overall scale, as opposed to each estimate being generated separately and the distribution happening to fit a logarithmic function. Under most circumstances, it is impossible to discriminate between these two possibilities. However, within a microgenetic study, that assesses representations on a trial-by-trial basis while children are receiving potentially instructive experience, it is possible to test whether the change is broad and sudden, as would be expected if one coherent representation is substituted for another, or whether the change is local and gradual, as would be expected if each estimate is generated solely on the basis of the particular number being estimated and feedback is interpreted in terms of its implications for the particular number and those near it.

The three feedback conditions that children were presented controlled for a great many alternatives to the log discrepancy hypothesis. The choice of numbers around 5 and around 725 as the comparison points for numbers around 150 was due in part to those ranges including one set of numbers smaller than 150 and one set larger than 150. Thus, if gains in the 150-feedback condition exceeded gains in either of the other feedback conditions, the effect could not be explained by anchoring pulling estimates up or down (because 5 would be a stronger anchor for pulling estimates down, and 725 would be a stronger anchor for pulling estimates up). The two conditions also controlled for the possibility that the key was for the feedback to pull estimates away from the extremes (725 is closer to 500 than is 150) and for the possibility that the key was for the feedback to indicate that extreme estimates were acceptable (5 being closer than 150 to an end of the number line). The numbers around 5 and 725 also were highly similar in the discrepancy between logarithmic and linear representations, thus allowing a test of whether this similar discrepancy would lead to similar learning.

The no-feedback control condition provided unique information, as well as serving as a point of comparison for learning in the other conditions. One type of information involved the stability of the assessments of logarithmic representations. If, in fact, a logarithmic pattern of estimates reflects a stable underlying logarithmic representation, subsequent estimates, in the absence of feedback, should also conform to a logarithmic pattern. No previous experiment had tested this prediction. The control condition also tested whether regression to the mean on those estimates that were most discrepant from their correct placement could account for changes in estimates in the three feedback conditions. The prediction was that in those three feedback conditions, there would be substantial changes in the maximally discrepant estimates (those around 150), but that this would not be the case in the control condition.

The design of Experiment 2 also allowed us to learn about five key dimensions of cognitive change: the source, rate, path, breadth, and variability of change. These dimensions have been proposed as central aspects of change within overlapping waves theory and have proved useful in describing cognitive change in a wide variety of contexts (for reviews, see Siegler, 1996, 2006). 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 items or whether the greatest change occurred on the numbers where the discrepancy between the logarithmic and linear representations was greatest, regardless of the distance of those items from the feedback items. Finally, to enhance understanding of the *variability* of change, we examined whether children whose pretest estimates adhered most closely to a logarithmic function learned more than other children and adhered more closely to the linear function on the posttest. The logic underlying this hypothesis was that children probably differed in the degree to which their estimates were based on their representation of the entire numerical range, as opposed to their knowledge of particular numbers, and that substitutions of one representation for another were most likely for children whose estimates were most influenced by the overall representation.

3.1. Method

3.1.1. Participants

The children in Experiment 2 were the 61 second graders (mean age = 8.2, $SD = 0.6$) whose estimates in Experiment 1 better fit a logarithmic than a linear function. A female research assistant served as experimenter.

3.1.2. Task

As in Experiment 1, each problem consisted of a 25 cm line, with the left end labeled "0," the right end labeled "1000,"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 close to the focal number for each feedback condition (5, 150, and 725); these 7 numbers ranged from 3 below the relevant focal number to 37 above it. The purpose was to include enough values to assess the local as well as the broad effects of each feedback condition.

3.1.3. Design and procedure

Children were randomly assigned to four experimental conditions: 150-feedback, 5-feedback, 725-feedback, or no-feedback. As shown in the outline of the procedure in Table 1, children in all four groups completed the number-line estimation task for three trial blocks and a posttest; administration of these trial blocks and feedback immediately followed Experiment 1. For children in the three feedback groups, each trial block included a feedback phase and a test phase. As shown in Table 1, the feedback phase of each trial block included either one item on which children received feedback (trial block 1) or three items on which they received feedback (trial blocks 2 and 3). The test phase in all three trial blocks included 10 items on which children did not receive feedback; this test phase occurred immediately after the feedback phase in each trial block. Children in the no-feedback group received the same number of estimation trials, but they never received feedback. On the posttest, children in all four groups were presented the same 22 problems without feedback as in Experiment 1. The children's estimates in Experiment 1 provided pretest data, which was used as a point of comparison for their subsequent performance and was obtained during the same session.

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

Table 1
Design of Experiment 2

Group	Phase						
	Trial block 1		Trial block 2		Trial block 3		Posttest
	Feedback[a] (1 item)	Test (10 items)	Feedback[a] (3 items)	Test (10 items)	Feedback[a] (3 items)	Test (10 items)	(22 items)
150-feedback	150	0–1000	147–187	0–1000	147–187	0–1000	0–1000
5-feedback	5	0–1000	2–42	0–1000	2–42	0–1000	0–1000
725-feedback	725	0–1000	723–763	0–1000	723–763	0–1000	0–1000
No-feedback	150, 5, or 725	0–1000	147–187, 2–42, or 723–763	0–1000	147–187, 2–42, or 723–763	0–1000	0–1000

[a] During the feedback phases, one-third of children in the no-feedback group were asked to estimate the positions of the same numbers as children in the 5-feedback group (but without feedback), one-third were asked to estimate the positions of the same numbers as children in the 150-feedback group (without feedback), and one-third were asked to estimate the positions of the same numbers as children in the 725-feedback group (without feedback).

the first trial block, to mark the positions of 3 numbers from 147 to 187 on the second trial block, and to mark the positions of a different 3 numbers from 147 to 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 to 42 on the second trial block, and of a different 3 numbers from 2 to 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 to 762 on the second trial block, and of a different 3 numbers from 722 to 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 on that page a 25 cm ruler (hidden from the child) that indicated the location of every 10th number from 0 to 1000. 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. For example, if a child were asked to mark the location for 18 (i.e., N) and his estimate corresponded to the actual location of 200 (i.e., N_{estimate}), the experimenter would write the number 200 above the child’s mark and mark where 18 would go on the number line. 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?”

3.2. Results and discussion

3.2.1. 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 Fig. 3, on the pretest, second graders’ median estimates for each number were better fit by the logarithmic function than by the linear one for children in all four groups. The precision of the fit of the logarithmic function, and the degree of superiority of that function to the linear function, was similar across the four conditions (5-feedback: log $R^2 = .95$, lin $R^2 = .71$, t [21] = 2.71, $p < .05$, $d = .80$; 150-feedback: log $R^2 = .95$, lin $R^2 = .72$, t [21] = 2.46, $p < .05$, $d = .73$; 725-feedback: log $R^2 = .93$, lin

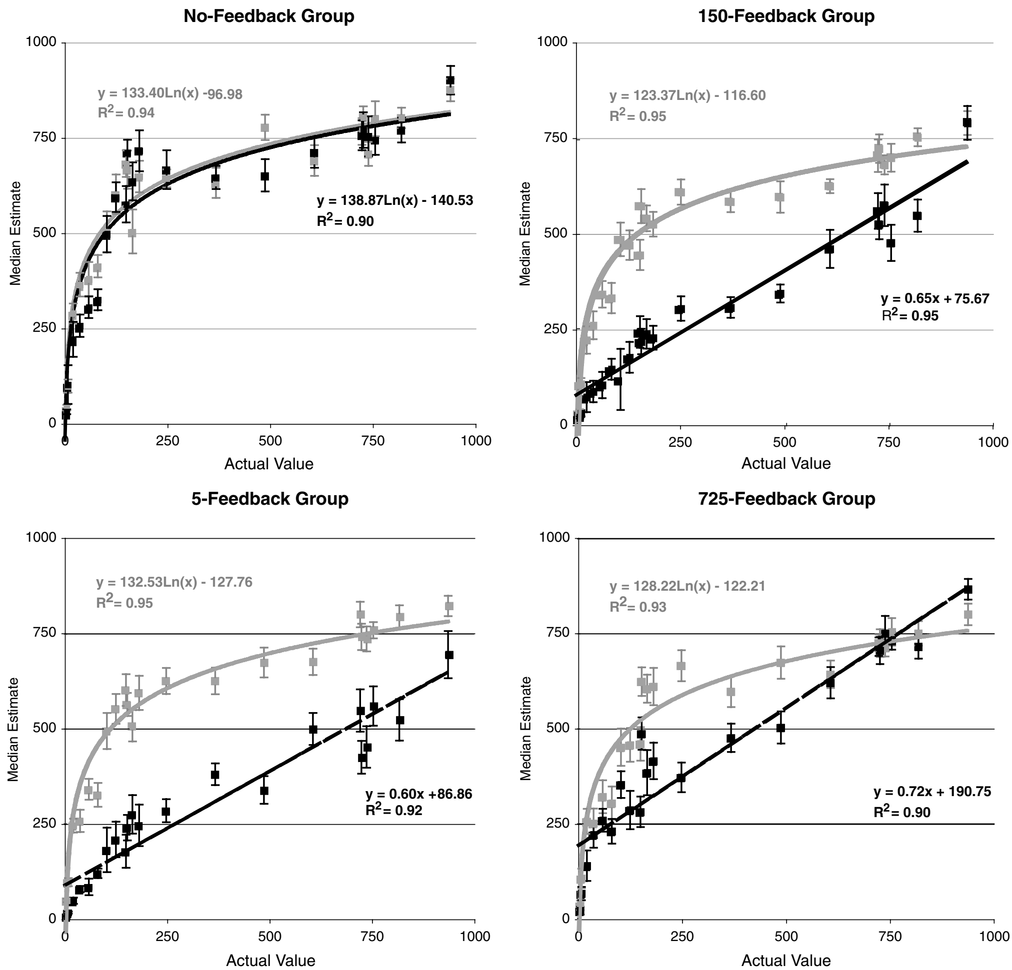

Fig. 3. Experiment 2. Best fitting functions for pretest (light colored) and posttest (dark color) median estimates. Solid function lines indicate that the function fit the data significantly better than the alternative model did. Dashed function lines indicate that the fit of the two functions did not differ significantly.

$R^2 = .68$, t [21] = 2.54, $p < .05$, $d = .73$; no-feedback: log $R^2 = .94$, lin $R^2 = .64$, t [21] = 2.39, $p < .05$, $d = .71$).

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, $p < .05$, $d = .78$). 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, n.s.; 725-feedback: lin $R^2 = .91$, log $R^2 = .84$, t [21] = 1.18, n.s.). 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, $p < .05$ $d = .66$). 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.

A comparison of percent absolute error on the pretest and posttest revealed that differences in improvements in accuracy among the four conditions were also present, $F(3, 87) = 2.82$, $p < .05$. On this measure, performance of children in all three feedback conditions improved more than performance of children in the no-feedback control d's $\geqslant 1.15$, with no significant differences among the feedback groups. This measure, however, may be less meaningful regarding changes in representations than it might appear. As can be seen in Fig. 3, children tended to lower all their estimates from pretest to posttest, thereby increasing accuracy partly by lowering estimates and partly by providing more linear estimates.

3.2.2. 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 (if at all) in the no-feedback group.

A 4 (training group: 5-, 150-, 725-, or no-feedback) × 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, $p < .001$, $d = 3.77$), 5-feedback group (29% of trial blocks, $p < .001$, $d = 1.33$) or 725-feedback group (37% of trial blocks, $p < .05$, $d = .86$). 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 $< .01$, d's $\geqslant 1.33$). 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 (Fig. 4) 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% by definition 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 ($p < .05$). What this meant was that the superiority of the 150-condition for promoting learning manifested itself after feedback on a single estimate. Providing feedback on the single number 150 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 after that one feedback problem ($p < .001$). 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 $< .05$).

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

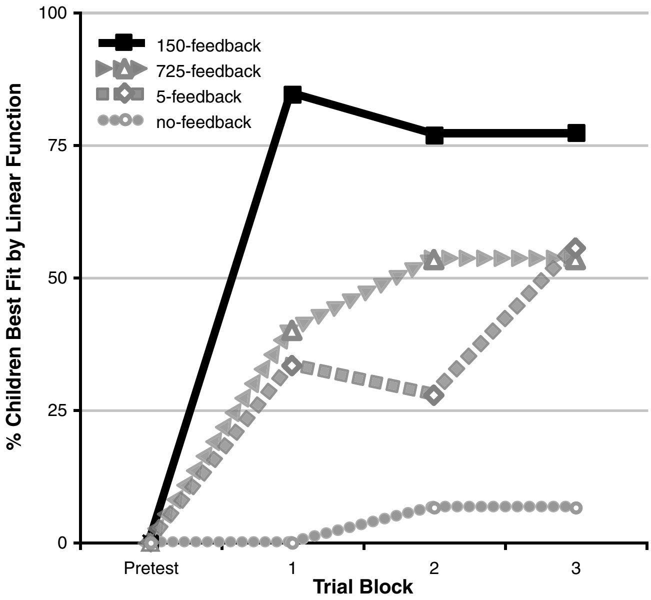

Fig. 4. Experiment 2: trial block-to-trial block changes in percentage of children in each condition whose estimates were best fit by the linear function.

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, $t\ (25) = 2.37$, $p < .05$, $d = .92$). 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 problem or problems that led to its initial adoption. Both the rapidity of the change in estimates and its stability once it was made suggest that the change was made at the level of the entire representation, rather than as a local repair, a conclusion that the data on the breadth of change also supported.

3.2.3. Breadth of change

To examine the breadth of change in children's estimates, we first examined the percentage of the 22 items on which mean absolute errors of children in the four groups were lower on the posttest than on the pretest.. 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 for that item. 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% improvement, respectively). The improvements for the numbers around 150 (147, 150, and 163) were noticeably larger (36%, 31%, and 25% 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% 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%). In the 150-feedback condition, both the generalization gradient and log discrepancy hypotheses led to the same prediction—that numbers around 150 should show especially large improvements (which they did, 21%, 30%, and 25% 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 for each number against the discrepancy between logarithmic and linear representations of that 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 for these functions used in Siegler and Opfer (2003). These equations were chosen so that both functions would pass through 1 and 1000.

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, $R^2 = .12$, $F(1, 21) = 2.84$, n.s.

To appreciate just how powerful the relationship was in the three feedback groups, 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. Particularly striking, children who received feedback on numbers from 722 to 762, like the other children, showed the greatest improvement on numbers from 56 to 246. 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 were in this range.

These results suggested three 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 the areas of the number line on which children received feedback; the improvements extended to quite distant areas. 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. Third, the change was not attributable to the estimates that were initially least accurate regressing to the mean level of accuracy; estimates of children in the control group did not improve on the same items. Analyses of children's path of change lent additional support to these conclusions, as described in the next section.

3.2.4. 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, suggested by incremental theories of representational change, was that the path of change entailed gradual, continuous improvements in the linearity of estimates. According to this hypothesis, the fit of the linear model would have gradually increased, and the fit of the 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 the point at which children first chose a different representation.

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 high 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. As shown in Fig. 5, 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 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 $R^2 = .80$, $F(1, 75) = 25.67$, $p < .001$, $d = 1.16$. 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 $R^2 = .64$, $F(1, 75) = 4.95$, $p < .05$, $d = .51$. 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.

3.2.5. 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

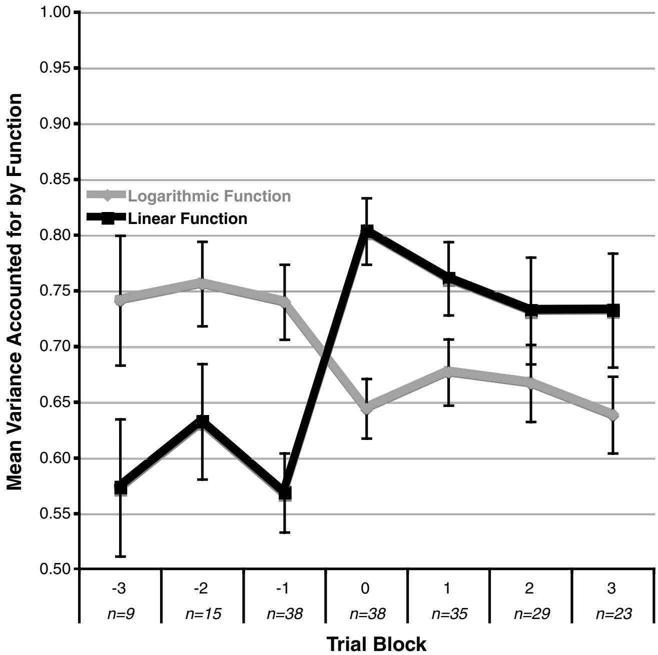

Fig. 5. Experiment 2: backward trials graph of fit of linear and logarithmic models to children's estimates. The 0 trial block is the block on which the linear function first provided a better fit to each child's estimates; the −1 trial block is the block before that, and so on. The *N*'s indicate the number of children who contributed data at each trial block; thus, 38 children used the linear representation on a least one trial block and therefore contributed data to trial block 0, 35 of these children had at least one trial block after this point and therefore contributed data to trial block 1, and so on.

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 in pretest estimates accounted for by the logarithmic function with percent variance in posttest estimates accounted for by the linear function. 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 ($r = .36$, $F(1, 45) = 6.75$, $p < .05$, $w^2 = .11$). 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 ($r = .42$, $F(1, 29) = 6.02$, $p < .05$, $w^2 = .14$); it was not found among children whose posttest estimates were not better fit by the linear function ($r = .08$, n.s.). 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 ($r = .19$, n.s.) or those who did not ($r = .32$, n.s.). 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 accuracy would be greatest.

4. General discussion

The present experiments yielded an unusually clear description of a cognitive transition. They also lent support to the depiction of representational change offered by the overlapping waves theory in general and the log discrepancy hypothesis in particular. In the remainder of this article, we summarize the present findings and discuss the broader implications of these findings for addressing three questions that are central within developmental, cognitive, and educational psychology: (1) how does change occur? (2) what cognitive processes can produce broad and rapid changes? and (3) what educational interventions will produce broad and rapid learning?

4.1. How does change occur?

Classic cognitive developmental theorists have proposed three main perspectives on representational change: stage theories, incremental theories, and early competence theories. Stage theories (e.g., Piaget, 1964; Bruner et al., 1966) depict representational development as involving broad and abrupt shifts from a less mature representation (e.g., enactive representations) to a more mature one (e.g., symbolic representations). Within this approach, children are initially limited to the less mature way of representing the information necessary to solve problems; improved problem solving occurs when children substitute the more advanced representation for the less advanced one. Incremental theories (e.g., Brainerd, 1978) also depict children as possessing only one representation at a time for a given type of problem. Such theories differ from stage-based approaches in that they depict representational development as occurring slowly and continuously, varying with the particular problem, and gradually spreading from problems on which children gain experience to increasingly dissimilar problems. Finally, early competence theories (e.g., Gelman & Gallistel, 1978), like stage and incremental theories, depict children as utilizing a single representation. Unlike the depiction in the other two theories, however, this single representation is present and embodies the key principles in the domain from early in development. Improved problem solving is produced by elaboration of the basic representation and improved understanding of task demands, rather than changes in the basic representation.

Overlapping waves theory differs from these classic approaches in that it depicts individual children as generally knowing and using multiple, co-existing representations. Within this theory, representational change is typically incremental; children gradually increase their reliance on more advanced representations, as well as occasionally adding new representations to the mix. However, in situations in which children are exposed to novel information that suggests that a representation that they use in other contexts yields much more accurate performance in a new, relatively similar, context, broad and abrupt representational change can occur (Siegler, 2006). The log discrepancy hypothesis was an attempt to specify a class of situations under which such atypically broad and abrupt representational change might occur, at least for nearby numerical ranges (Dowker, 2003).

The present results suggest that it may be useful to broaden the log representation hypothesis into a general *representational discrepancy hypothesis*: broad and abrupt representational change is likely in situations in which children (1) initially rely on an immature representation, (2) rely on a more mature representation in another, structurally parallel domain, and (3) have experiences that highlight discrepancies between the less and more mature representations and make clear the superiority of the more mature one. Testing the representational discrepancy hypothesis will require meeting the three conditions in other domains and determining whether similarly rapid and broad representational change occurs in them. In at least one context other than the present one in which the three conditions were met, Opfer and Siegler's (2004) study of the concept "living things," the predicted broad and abrupt representational change did occur. Once children gained a more mature representation of the kinds of things (plants and animals) that can move adaptively (e.g., turn toward food or sunlight), they used the new approach consistently and with items as diverse as grass, trees, and bears.

Alternative perspectives on representational change may be evaluated by their ability to predict findings regarding the five dimensions of change that we examined in the present study: source, rate, path, breadth, and variability. Our findings regarding each of these dimensions of change can be summarized quite succinctly. Feedback on numerical magnitudes was a potent source of change, especially feedback on magnitudes around 150, the range in which logarithmic and linear representations of the 0–1000 range are maximally discrepant. The rate of change was very high, with children often switching from a logarithmic to a linear representation after a single feedback trial. The path of change involved a direct transition from a well fitting logarithmic pattern of estimates to an equally well fitting linear pattern; there were no intermediate forms, nor much oscillation between the two representations. The change was broad, encompassing the entire numerical range from 0 to 1000, with the largest change occurring not on problems where feedback was given but rather on problems where the logarithmic and linear representations were most discrepant. Finally, variability among children in initial adherence to the logarithmic representation was positively related to subsequent learning; the closer the fit of a child's pretest estimates to the logarithmic function, the closer the fit of the child's posttest estimates to the linear function.

These findings provide a clear description of the changes in numerical estimation observed in the present study. The more difficult challenge, as always, was to specify *how* the changes occurred. Data from the present and previous experiments on number line estimation suggested the following account: (1) second graders initially represented the 0–1000 range logarithmically; (2) they already represented smaller numerical ranges, such as 0–100, linearly; (3) feedback led to substantial and rapid changes in estimates, with the quickest and largest changes occurring when the discrepancy between the logarithmic and linear representations was greatest; (4) the change occurred at the level of the entire representation, rather than in a more local and piecemeal way; (5) analogical mapping from smaller to larger numerical contexts was the key change mechanism. In the remainder of this section, we lay out the evidence for the first four points within this account in greater depth, and evaluate general theories of representational development in light of this evidence. In the next section, we describe the fifth, more speculative point within the account, and consider its implications for mechanisms of cognitive change.

4.1.1. Second graders initially represented the 0–1000 range logarithmically

Siegler and Opfer (2003) found that second graders' number line estimates fit a logarithmic pattern, and hypothesized that children of this age used a logarithmic representation of numerical magnitudes to generate these estimates. They also found that many fourth graders and almost all sixth graders generated patterns of estimates that fit a linear function, and therefore hypothesized that with age and experience, children move from a logarithmic to a linear representation of numerical magnitudes in this range.

The present findings supported this conclusion. As predicted, the improvement in estimation accuracy between second and fourth grade was greatest for numbers around 150, the range where logarithmic and linear representations differed by the largest amount. This range was not included in Siegler and Opfer's stimulus set, and it is unclear what perspective other than that of movement from a logarithmic to a linear representation would have predicted the finding. Also, the logarithmic function fit the pretest estimates of two-thirds of second graders in the Experiment 1 sample, and these were the children who participated in Experiment 2. Thus, rather than children starting with a correct, linear representation of numerical quantities (as predicted by early competence theories), there is a substantial period of time during which children have an incorrect understanding of the correspondence between numerals and magnitudes, leading to reliance on a default logarithmic representation that is widely used across species and situations (Dehaene, 1997).

4.1.2. The children had available a linear representation of smaller numerical magnitudes

In Siegler and Opfer (2003), the same second graders who generated logarithmic patterns of estimates on 0–1000 lines often generated linear patterns on 0–100 lines. Both Siegler and Booth (2004) and Booth and Siegler (2006) also found that most second graders generated linear patterns of estimates on 0–100 number lines. These data, together with the Experiment 1 data from the present study, made it likely that most second graders in the present experiment also had available the linear representation for the 0–100 range. If children did not already know and use the linear representation in that range, such a rapid shift in the 0–1000 range would have been unlikely.

4.1.3. The larger the discrepancy between the two representations for the numbers where feedback was provided, the more often the feedback elicited substantial change

Several types of evidence suggested that the discrepancy between the logarithmic and linear representations, and therefore between children's estimates and the correct values, was a key factor in determining the effects of feedback on learning. The strongest supporting evidence was that the change stimulated by the 150-feedback condition was more rapid and substantial than that in the other two feedback conditions or in the no-feedback condition. This was the numerical range where the logarithmic and linear functions were most discrepant. A second type of evidence for the importance of the discrepancy in triggering the change came from the individual differences data. Children whose pretest estimates fit the logarithmic function most precisely, and who therefore had the largest discrepancies between their estimates and the correct values, actually fit the correct (linear) function most closely on the posttest. Thus, these children's changes in estimates from pretest to posttest were also the largest. Sometimes, having a clear but wrong idea, and adhering to it systematically, may be more conducive to learning than having vague ideas or not having any clear idea. (At other times, having vague or conflicting ideas is more conducive to change, see Graham et al., 1993).

4.1.4. The change occurred at the level of the entire representation

Many, probably most, cognitive developmental changes are narrow in scope and are made in a piecemeal rather than an integrated fashion. For example, learning of past tense verb forms for irregular verbs occurs on a word-by-word basis, with "ed" overgeneralization errors disappearing very early on some verbs but only years later on others (Marcus et al., 1992). Similarly, preschoolers who learn that $4 + 2 = 6$ do not necessarily infer that

$2 + 4 = 6$ (Geary, 1994), manual gestures on a problem often show increased understanding following experiences that leave verbal statements on the same problem unchanged (Albali, 1999; Goldin-Meadow, 2001), and so on.

The change observed in the present study was different; it seemed to occur at the level of the entire representation, rather than in a piecemeal fashion. Perhaps the most striking evidence that the change occurred at the level of the entire representation was that estimates improved most dramatically not on the problems on which children received feedback but rather on the problems on which the discrepancy between the two representations was largest. In the most extreme case, estimation accuracy of children in the 725-feedback group improved from pretest to posttest on all numbers below 500, despite those numbers being far from the numbers on which children in the group received feedback. Although performance in this range was less than perfect, this pattern seems extremely unlikely if change did not occur at the level of the representation as a whole. On the other hand, the result is directly predicted if children substituted a linear for a logarithmic representation on the task.

The rapidity of the change also pointed to the change occurring at the level of the entire representation. After a single feedback trial, the best fitting function switched from logarithmic to linear for 85% of children in the 150-feedback condition and for more than half of all children who received feedback. Moreover, once children's estimates first conformed to the linear function, the linear model continued to provide the best fit on more than 80% of subsequent trial blocks, regardless of the problems that led to the apparent switch of representations.

Data from Booth (2005) suggest that changes in representations of numerical magnitude, once made, are stable over time and general across a range of tasks. In particular, first graders who saw displayed linear representations of the magnitudes of addends of novel addition problems not only learned more addition facts at the time but continued to show greater knowledge two weeks later. Presentation of the linear representations of the addends also led to improved number line estimates. These findings lend additional support to the view that changes in numerical representations can occur at the level of the entire representation, rather than as incremental repairs to the original representation.

4.2. What cognitive processes can produce broad and rapid changes?

Although this conclusion is more speculative than the other aspects of the account of the change process, several types of evidence suggest that analogical mapping of an existing linear representation onto a new numerical range was an important mechanism of change. First, the analogy was there to be drawn and was relatively straightforward. The decimal system allows a direct mapping between the 0–100 and 0–1000 ranges, and the linear representation that most second graders use on the 0–100 number line task is directly applicable to the 0–1000 task. A second relevant type of evidence was the speed with which children's estimates, and presumably their representations of the numerical range, changed. Making such far reaching changes after a single feedback trial seems more likely if children were mapping an existing representation onto a new range of numbers than if they were constructing a totally new representation. A third relevant type of evidence involves the widespread importance of analogical mapping in other contexts (Chen & Klahr, 1999; Gentner, Holyoak, & Kokinov, 2001; Holyoak & Thagard, 1995; Opfer & Siegler, 2004). Both children and adults often solve problems and learn through drawing analogies to more familiar and better-understood situations, particularly where the mapping between the familiar and new situations is straightforward, as it was in the present situation. Given that the change of representations often occurred in a single trial, symbolic models of analogical reasoning (e.g., Gentner, Rattermann, Markman, & Kotovsky, 1995) or hybrid models (e.g., Hummel & Holyoak, 2003) seem particularly promising as a means for specifying the change process in greater detail and for evaluating whether the structured representations that we hypothesize to be central to rapid and abrupt change in numerical estimation could be accomplished in ways that differ substantially from the present account.

This account can and should be tested further empirically. If analogical mapping is the key change mechanism, we would expect children to have greater difficulty extending the linear representation to numerical contexts in which the mapping is less straightforward than in the present context. For example, second graders might have more difficulty mapping the linear representation from 0–100 to 8–108 or 18–82, despite the individual numbers overlapping more with the 0–100 range on which the children already use the linear representation. Providing feedback on 8–108 or 18–82 number line problems would be expected to elicit weaker, slower, and narrower change than that observed in the present study of performance on 0–1000 number lines because children would have more difficulty drawing the relevant analogy. In contrast, because of the greater overlap of numbers in the 8–108 and 18–82 range with those in the 0–100 range, an associative or PDP-based account would likely make the opposite prediction.

4.3. What interventions will induce rapid and broad changes?

What educational interventions might induce rapid representational changes like the ones observed in the present study? With regard to number line estimation, the effectiveness of an intervention seems likely to depend partly on the transparency of the relation between the base and target—i.e., that the analogy be relatively straightforward—but also on what learners bring to the learning environment, including knowledge of the ordering of the numbers involved, of a base case from which they can analogize, and also perhaps evolved predispositions to represent numeral magnitudes in particular ways (Geary, 1995). If any of these types of knowledge is missing, the gap would have to be remedied before adoption of a new representation that produced broad and rapid change would be likely.

Looking beyond number line estimation, we can ask: what are the features of interventions that make rapid changes of representations likely in general? The same three types of variables seem likely to be crucial. First, learners need to have a basic understanding of the elements within the target case. For example, the types of feedback that produced large changes in the present context would almost certainly not have produced comparably large and rapid changes if the situation involved fifth graders being asked to place decimal fractions with varying numbers of digits on a number line. Children of this age rarely order decimal fractions correctly (Rittle-Johnson, Siegler, & Alibali, 2001), and without such correct ordering of the magnitudes of the stimuli in the domain, the feedback would be extremely difficult to interpret. Second, the mapping between the base and target cases must be transparent. People often fail to draw useful analogies, and children are especially unlikely to draw the intended analogy, when it is unsupported by perceptual cues or when perceptual cues point to irrelevant analogies (Gentner, 1983; Holyoak, Junn, & Billman, 1984). Third, learners need to have strong knowledge of a base case from which they can analogize. Lack of deep understanding of the casual and structural relations within the base case can prevent drawing the ideal analogy, even when the other two elements are in place.

When all three conditions are met, however, rapid changes in representations can occur in domains quite different from number line estimation. This was evident in a recent study of the concept of living things (Opfer & Siegler, 2004). The 5-year-olds in that study had considerable knowledge of the relevant elements within the target category—plants—but did not believe that they are living things. However, the 5-year-olds did have strong knowledge of a base case from which they could analogize, namely animals. After learning a relevant fact—that plants can move in ways that promote their survival—most 5-year-olds inferred that plants are living things as well. As in the present context, the children seemed to change their representations on the basis of an analogy, in this case reasoning that if plants, like animals, move in ways that are important for their survival, then plants, like animals, are living things. The change in behavior following the change in representations of the living things concept was both broad and rapid. Judgments of life status underwent large and rapid changes for such diverse category members as grass, trees, bushes, and flowers; the changes occurred within one or a few feedback trials. Determining more precisely how to produce such broad and rapid change in understanding seems important not only for improving educational practice but for enhancing theoretical understanding of representational change as well.

References

Albali, M. W. (1999). How children change their minds: strategy change can be gradual or abrupt. *Developmental Psychology, 35*, 127–145.

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

Booth, J. L. (2005). *The importance of an accurate understanding of numerical magnitudes*. Unpublished doctoral dissertation, Carnegie Mellon University, Pittsburgh, PA.

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

Brainerd, C. J. (1978). The stage question in cognitive developmental theory. *Behavioral and Brain Sciences, 1*, 173–213.

Bruner, J. S., Olver, R. R., & Greenfield, P. M. (1966). *Studies in cognitive growth*. New York: Wiley.

Chen, Z., & Klahr, D. (1999). All other things being equal: children's acquisition of the Control of Variables Strategy. *Child Development, 70*, 1098–1120.

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.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. NY: Oxford University Press.

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

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

Geary, D. C. (1994). *Children's mathematical development: Research and practical implications*. Washington, DC: American Psychological Association.

Geary, D. C. (1995). Reflections of evolution and culture in children's cognition: Implications for mathematical development and instruction. *American Psychologist, 50*, 24–37.

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

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

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

Gentner, D., Rattermann, M. J., Markman, A., & Kotovsky, L. (1995). Two forces in the development of relational similarity. In T. J. Simon & G. S. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 263–313). Hillsdale, NJ: Erlbaum.

Goldin-Meadow, S. (2001). Giving the mind a hand: The role of gesture in cognitive change. In J. L. McClelland & R. S. Siegler (Eds.), *Mechanisms of cognitive development: Behavioral and neural perspectives* (pp. 5–31). Mahwah, NJ: Erlbaum.

Graham, T., & Perry, M. (1993). Indexing transitional knowledge. *Developmental Psychology, 29*, 779–788.

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

Holyoak, K. J. (1978). Comparative judgments with numerical reference points. *Cognitive Psychology, 10*, 479–486.

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

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

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

Laski, E., & Siegler, R. S. (2005). *Children's number categories and their understanding of numerical magnitude*. Unpublished manuscript.

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.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., & Xu, F. (1992). Over-regularization in language acquisition. *Monographs of the Society for Research in Child Development, 57*(4 Serial No. 228).

Moyer, R. S. (1973). Comparing objects in memory: Evidence suggesting an internal psychophysics. *Perception and Psychophysics, 13*, 180–184.

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

NCTM (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

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

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

Piaget, J. (1964). Development and learning. In T. Ripple & V. Rockcastle (Eds.), *Piaget rediscovered* (pp. 7–19). Ithaca, NY: Cornell University Press.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: an iterative process. *Journal of Educational Psychology, 93*, 346–362.

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

Siegler, R. S. (2006). Microgenetic analyses of learning. In W. Damon & R. M. Lerner (Series Eds.) & D. Kuhn & R. S. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2: Cognition, Perception and Language*. (6th ed., pp. 464–510). Hoboken, NJ: Wiley.

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.

Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *The origins of cognitive skills* (pp. 229–293). Hillsdale, NJ: Erlbaum.

Sowder, J., & Wheeler, M. (1989). The development of concepts and strategies used in computational estimation. *Journal for Research in Mathematics Education, 20*, 130–146.