COGNITION AND INSTRUCTION, 1990, 7(1), 55–78

Development of Numberline and Measurement Concepts

Andrea L. Petitto
Graduate School of Education and Human Development
University of Rochester

In a two-part study, children in Grades 1 through 3 were given a set of numberline estimation problems and a set of similar numberline estimation problems that were to be solved using specially designed rulers. The estimation tasks showed a complex pattern of change in accuracy and strategies over the three grade levels. This pattern indicated that children progressively altered their estimation strategies from strictly sequential ones to ones that incorporate elements of proportional reasoning. This shift was gradual and implemented school-learned numerical relationships in the solution of these novel tasks. There was no indication that, in the Piagetian sense, stage changes contributed to strategy development. However, the use of rulers with different spacings between the numbers indicated that children's acquisition of a principled concept of equal intervals required both the learned convention of equal intervals and the achievement of conservation of length.

The two experiments presented in this article were designed to explore the development of the related concepts of numerical proportion and units of measurement. The impetus for these experiments with paper-and-pencil representations of numberlines and rulers was provided by earlier observations of elementary school children playing a computerized numberline estimation game. When making successive approximations to a precise target, the children used many different strategies, some of them quite puzzling to an adult observer. For example, children occasionally persisted in using small incremental adjustments to an original estimate, even when that original estimate was visibly far off the mark (Levine, Boruta, & Petitto, 1983). This led to questions about the young child's conceptualization of the spacing and proportion relations of numbers on a numberline or ruler.

These experiments examine two factors that contribute to a child's devel-

Requests for reprints should be sent to Andrea L. Petitto, Graduate School of Education and Human Development, University of Rochester, Rochester, NY 14627.

oping theory of measurement in the context of the linear ruler: the acquisition of specific skills in counting as well as arithmetic and conceptual differentiations in the child's understanding of linear space. The elementary school curriculum is both a source of direct instruction about the structure and function of rulers and a source of component skills in counting and arithmetic that can be applied in the ruler context. Changes in the ways that children conceptualize and manipulate linear space that are not directly related to the school curriculum could also affect their ability to organize curricular elements into a coherent theory of measurement.

The way young children map numbers onto a numberline appears to reflect, in part, their understanding of cardinal and ordinal relationships (Gelman & Gallistel, 1978), but does not necessarily reflect an understanding of proportional relationships. Young children can verbally count forward in one direction along a ruler or numberline[1] and count backward in the opposite direction. Although this counting procedure suggests an appreciation of the mapping of counting sequences, the proportion relationships that specify the distances between numbers might not be similarly appreciated. Fuson (1984) observed that, in the early elementary years, children rarely understand that "Numbers are represented on the numberline by lengths—instead, numbers are thought to be represented by the points they label" (p. 219). How do children progress from a strictly sequential concept of numerical relationships to one that incorporates proportional relationships?

Part of the answer might be in the acquisition of arithmetic and counting skills. These skills may contribute to a mental representation of numerical relationships that specify proportional relationships among numbers. For example, knowing numerical doubles and halves might make it easier for children to translate that relationship to a spatial numerical scale such as a linear ruler. Recent studies have shown that children's representation of numbers becomes more complex and begins to incorporate arithmetic relationships as the children pass through the elementary grades (Miller & Gelman, 1983; Murray & Mayer, 1988). The work of Ashcraft and his colleagues (Ashcraft & Battaglia, 1978; Ashcraft & Fierman, 1982; Ashcraft & Stazyk, 1981) suggests that one source of this increasing complexity is the acquisition of arithmetic knowledge, particularly multiplication. Numerical proportions specified by sums and multiples might then be mapped onto corresponding relationships in a spatial representation such as a numberline or ruler. Instead of perceiving the numbers on a ruler as simply an ordered

[1]Because this research concerns the relationships between numerical values and the subdivision of linear space rather than measurement standards such as inches and centimeters, there is little distinction between the ruler concept and the concept of the integer numberline as they are used here. Accordingly, the term *ruler* is used when referring to a physical object, and the term *numberline* is used when referring to a generalized concept of linear spatial layout, of which rulers are one physical manifestation.

sequence, the spaces between the numbers would be seen as corresponding to the additive and multiplicative properties of the numbers themselves; the distance between 0 and 1 would match the distance between 1 and 2, because 1 + 1 = 2, and so on.

Changes in the reasoning process about spatial extent might also be a factor in the child's acquisition of measurement concepts. Piaget, Inhelder, and Szeminska (1960) postulated the necessity for several generalized cognitive operations (e.g., conservation of length), indicating that a child recognizes the invariance of length over changes in position. The acquisition of *conservation of length* requires the child to move from an initial conflation of the concepts of position and length to a more differentiated conceptual structure that distinguishes clearly between position and length. It is reasonable to expect that such a conceptual advance might be important in the transition from seeing numbers on the numberline as a sequence of positions to understanding that those numbers relate to a progression of lengths.

The two experiments presented in this article explore the influence of component arithmetic and counting skills and of changes in children's conceptualization of linear space and its measurement. The first experiment demonstrates that children's conceptualization of the layout of numbers on a numberline undergoes qualitative change from Grades 1 through 3 and assesses the relative contributions of the acquisition of component skills and of conservation of length to this qualitative change.

The second experiment examines the development of a principled concept of equal-interval subdivisions on a ruler by testing the existence and strength of children's belief that linear measures must be subdivided into equal intervals. These developments are related to elements in the curriculum and to conservation of length.

EXPERIMENT 1

Using a task in which children were asked to estimate the numerical value of positions on a numberline, this experiment was designed (a) to establish the nature of qualitative changes in children's conceptualization of the layout of numbers on a numberline and (b) to assess the effects of acquiring component arithmetic skills and conservation of length on those qualitative changes.

A *sequence-to-proportion shift hypothesis* states that, through the first three grades, children shift from sequential- to proportion-based strategies for estimating the position values on numberlines. Sequential estimation strategies neglect the matter of distance between positions and take into account only one endpoint value when estimating the value of any other position on the line. Proportion-based strategies relate numerical proportions to spatial proportions on the numberline. Such strategies include the identification of positions near the midpoint, with values intermediate between the two endpoint values and the use of alternative counting intervals, such as 1s, 5s, and 10s, to accommodate different numerical ranges indicated by endpoints.

Two hypotheses concern the role of classroom learning and the development of Piagetian operations in the expected sequence-to-proportion shift. The *component skills hypothesis* claims that the acquisition of counting and arithmetic skills is causally related to the observed qualitative changes in strategy and accuracy. If all that is needed are explicit skills in counting and arithmetic, then children should be able to apply those skills in appropriate ways shortly after they acquire them in school; the sequence of changes in children's strategies will closely follow the sequence of counting and arithmetic skills presented in school.

The *conceptual differentiation hypothesis* relates expected strategy shifts to the achievement of conservation of length. The ability to conceptualize the distinction between position and length could affect the ways that children apply their numerical knowledge to numberline representations. Even when children can count verbally by 5s and 10s, as well as by 1s, the decision to use interval counting and the spacing of the count across a line involves the comprehension of the subdivision of the line into lengths, as opposed to a simple sequence of positions.

Evidence for the sequence-to-proportion shift, and for the two hypotheses that explain the reasons for this shift, will be derived from patterns of accuracy and from observations of the children's overt strategies when solving numberline estimation problems. The sequence-to-proportion shift predicts a shift in counting strategies from counting by 1s from a single endpoint toward the target to using different counting intervals (e.g., 1s, 5s, or 10s) to accommodate different numerical ranges. It also predicts that, within the age range tested, children will begin to use familiar arithmetic proportions, such as doubles and halves, to identify the value of corresponding line positions such as midpoints.

The component skills hypothesis suggests the sequential incorporation of new strategies subsequent to their introduction in the curriculum. The conceptual differentiation hypothesis predicts that strategies based on these counting and arithmetic skills will only be effectively incorporated after the acquisition of conservation of length. If this conceptual differentiation is not necessary, new strategies should appear gradually, following the sequence in which they appear in the curriculum.

Strategies are not always observable, however, and patterns of accuracy of estimates can also be used to infer characteristics of reasoning. Based on the deviation of the children's estimates from exact target values, the hypothesized sequence-to-proportion shift predicts changes in the pattern of deviation scores for targets at different positions on the numberline and for numberlines with different numerical ranges. When children use sequential strategies, the deviation of their estimate from the exact value of the target would increase for targets further from the endpoints and, therefore, would be greatest for targets nearer the midpoint of a line. If children further limit their sequential strategies by orienting only to the 0 endpoint, the magnitude of error would increase for targets distant from the 0 endpoint so that deviations from the exact value would be greatest for targets near the high end of the line.

Deviations from the exact target values can also provide evidence for the component skills and conceptual differentiation hypotheses. An interaction between grade level and numerical range would support a component skills hypothesis. If young children use smaller, more familiar numbers, the magnitude of deviations from the exact target values would be exaggerated in higher numerical ranges, especially for targets near the middle or high end of the line. Differential improvement on high and low numerical ranges would result from increasing familiarity with larger numbers. The conceptual differentiation hypothesis would be supported by an abrupt change in deviation scores in both numerical ranges at about the same time and a significant correlation with conservation of length, whereas a gradual change over time would argue against conceptual differentiation as an important factor.

Method

Subjects. Eighty-one children from one elementary school in a suburban school district participated. There were 16 children (7 girls, 9 boys) from one first-grade classroom, 44 children (21 girls, 23 boys) from two second-grade classrooms, and 21 children (13 girls, 8 boys) from one third-grade classroom. Half of the second graders from each classroom were tested in the fall, and the remaining half were tested in the spring. All first and third graders were tested in the spring. This testing schedule procedure resulted in four distinct groups according to grade: late first, early second, late second, and late third. These groupings were intended to capture transitions between critical topics in the school curriculum, as well as the age range important for developmental changes in cognitive processes.

Table 1 shows the sequence of relevant topics covered in the school's elementary mathematics curriculum. The curriculum was designed by a committee of teachers in the district and was based, in part, on the New York State guidelines for elementary mathematics (New York State Education Department, 1985). Notice that the units on measurement include construction and use of both standard and nonstandard rulers from first grade onward. Rulers are used for measurement and are constructed by the iterative application of a single unit.

From first grade onward, the units on counting include counting by 1s,

TABLE 1
Outline of Pertinent Aspects of the Mathematics Curriculum

	Grade 1	*Grade 2*	*Grade 3*
Counting			
cardinal numbers	1s to 100 10s to 100 5s to 50	1s to 100 10s to 100 2s to 20 5s to 50	2s to 100 5s, 10s to 200 Rounding to nearest 10
Arithmetic			
Addition	Sums to 10 2 columns without regrouping	Sums to 18 2 & 3 columns with regrouping	3 & 4 columns with regrouping
Subtraction	Differences from 10 2 columns without regrouping	Differences from 18 2 & 3 columns, no regrouping Inverse relation to addition	3 & 4 columns with regrouping Inverse relation to addition
Multiplication		Multiples of 2 and 5 Relate multiplication and addition	Multiplication facts through 5 × 10 Relate multiplication and addition
Division		Halves of numbers	Divison facts through 5 Inverse relation of multiplication and division Relate division to subtraction fractions Divide whole objects into equal parts
Estimation	Estimate prior to counting and measuring	Estimate prior to counting and measuring	Estimate before measuring lengths
Measurement			
Nonstandard	Construct and use rulers from units	Repetitive use of a unit Construct and use rulers from units	Repetitive use of a unit
Standard	Measure in cm, dm, and m Introduce in. and ft	Measure in cm, dm, and m Introduce in. and ft	Repetitive use of cm unit Construct and use metric rulers

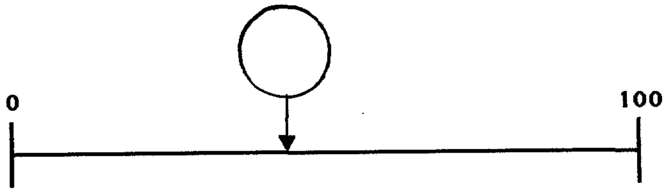

FIGURE 1 Example of a 0-to-100 range numberline problem.

5s, and 10s. Comments by the teachers indicated that even the weakest math students are able to count fluently by 10s by the beginning of second grade. Thus, the component skills hypothesis predicts the early use of different counting intervals (i.e., by 5s, 10s, or even 1s) by the beginning of second grade. Table 1 also shows that children learn doubles and halves during second grade.

Procedure. Children were given numberline estimation problems consisting of horizontal lines marked by numbers at the endpoints and by an unlabeled "target" arrow at some other point on the line (e.g., see Figure 1). The children were to determine what number would fall at the target point "if all the numbers were written in where they belong on the line."

Children were tested individually. Each child was given a conservation of length task first and then six numberline estimation problems. In the conservation of length task, two 10-in. sticks were presented to the child, lined up horizontally so that the ends matched. The child was then asked if the two were the same length or if one was longer. (All children said they were the same length.) As the child watched, the tester slid one stick about 2 in. to the right and repeated the question. This time, children were asked to justify their answer: "How do you know?" When children maintained that the two sticks were the same, their justifications were recorded and categorized as *identity* (logical necessity, e.g., "You didn't change anything; all you did was push it"); *compensation* ("This end sticks out as much as that end goes in"); *reversibility* ("If you push them back, they'll be the same"); or *no justification* (e.g., "I don't know"). Children were assigned to one of two levels of conservation based on their responses: "conservers," those who said the sticks were the same length and used any relevant justification and "nonconservers" if they said that the sticks were different lengths. No children fell into an "intermediate" stage, giving a conserving response with no justification.

The six numberline estimation problems were of the type shown in Figure 1. Three of these problems were marked 0-to-10 (0-to-10 range) and three others were marked 0-to-100 (0-to-100 range). The target position was different for each of the three problems in each numerical range. In the 0-to-10 range, the targets were at 2, 4, and 8; in the 0-to-100 range, the targets were at 16, 44, and 78. The following oral instructions were given to each child at the beginning of the set of six numberline problems:

> A numberline is like a ruler. It has numbers and marks all along it. This (indicating the first numberline problem) is a numberline too, but on this numberline only the numbers at the end points are written in. What are these numbers (indicating the marked end points and waiting while the child read them)? If all the numbers from zero to ten (zero to one-hundred) were written in where they belong, what number would be at the arrow? Write that number in the circle.

The children's strategies were observed and scored according to a coding scheme described in the Results section.

Children were randomly assigned to each of two orders of presentation: 10s first or 100s first. In each condition, all problems of that range were presented before all problems of the other range. The sequence of target positions was randomized across subjects, with the restriction that the sequence for each child was the same for both numerical ranges; that is, if the left, then right, then midtarget sequence was used in one numerical range, it was also used in the other range.

Both accuracy and strategy were observed in the children's performance. *Accuracy* is the difference between a child's estimate and the correct value at each target position; *strategies* are the observable components of the children's problem-solving activity. Relationships between changes in accuracy, the appearance of specific strategies, and the presentation of related skills in school were used to infer the processes children use to map numerical relationships onto the numberlines.

Results and Discussion

Conservation. Table 2 depicts the percentage of children who were conservers and nonconservers at each of the four grade levels. A chi-square analysis on the number of conservers in each grade showed a nonsignificant trend, $\chi^2(3, N = 81) = 7.08$, $p < .07$, for differences between grade levels. However, pairwise comparisons showed only the difference between late-second and late-third graders to be significant, $\chi^2(1, N = 43) = 6.24$, $p < .01$. The apparent drop in the number of conservers between late-first

TABLE 2
Percentage of Children Who Conserved Length

	Grade Level			
	Late 1st	*Early 2nd*	*Late 2nd*	*Late 3rd*
Conservers	56.25	54.55	50.00	85.71

TABLE 3
Averages and Standard Deviations of Deviation Scores for Each Grade Level at Each Target Position and Numerical Range (Average Above Standard Deviations)

	Grade Level				
	Late 1st[a]	*Early 2nd*[b]	*Late 2nd*[c]	*Late 3rd*[d]	*Overall*[e]
Range Position (0–10)					
Leftmost	15.38	13.81	11.50	5.00	11.08
	10.50	13.22	10.89	6.07	10.18
Midmost	17.69	22.86	18.50	3.00	15.41
	13.63	16.48	16.31	4.70	12.75
Rightmost	9.23	6.19	6.00	5.50	6.49
	13.20	9.73	5.98	6.86	8.55
Overall	14.10	14.29	12.00	4.50	10.99
	12.44	13.14	11.06	5.88	10.49
Range Position (0–100)					
Leftmost	11.15	10.19	8.25	6.30	8.78
	3.48	3.41	4.40	3.37	3.68
Midmost	25.38	28.38	18.70	9.65	20.17
	16.42	12.07	13.29	10.20	12.66
Rightmost	21.00	31.71	13.65	7.25	18.34
	17.77	26.64	14.50	5.45	16.07
Overall	19.18	23.43	13.53	7.73	15.76
	12.56	14.04	10.73	6.34	10.80

[a]$n = 13$. [b]$n = 21$. [c]$n = 20$. [d]$n = 20$. [e]$n = 74$.

and late-second graders was not significant. These results indicate little change in the number of conservers from late-first to late-second graders, but a sharp increase from late second to late third.

Deviation scores. Children's numerical solutions to the estimation problems were used to derive percent deviation scores according to the following formula:

$$\frac{|E-A|}{R} \times 100$$

in which E is the child's estimate, A is the actual target value, and R is the numerical range (10 or 100). This formula gives a comparable percent deviation across both numerical ranges. Table 3 summarizes these deviation scores for each grade level at each target position and numerical range, and the same scores are depicted in Figures 2 and 3. The three target positions in each numerical range are termed "leftmost" (2 or 16), "midmost" (4 or 44), and "rightmost" (8 or 78).

Figures 2 and 3 show that deviation scores are generally lower for 0-to-10 range than for 0-to-100 range. For each numerical range, there is considerable discrepancy in deviation scores at different target positions for early grades, but this discrepancy disappears by late-third grade. For each range, deviation scores on the target furthest from any endpoint (i.e., the mid-target) are high at early grade levels but drop from early-to-late second grade and from late-second to late-third grade. This pattern is consistent with the prediction of the sequence-to-proportion shift hypothesis, which posits that targets furthest from any endpoint would be more difficult than those near a marked endpoint.

The deviation score patterns for the leftmost and rightmost targets differ between the two numerical ranges. In the 0-to-10 range, scores on the leftmost target are higher than those for the rightmost target in early grades, but this pattern is reversed in the 0-to-100 range problems. This unexpected discrepancy between numerical ranges will be discussed in the following analyses.

The deviation scores were used as the dependent variable in an analysis

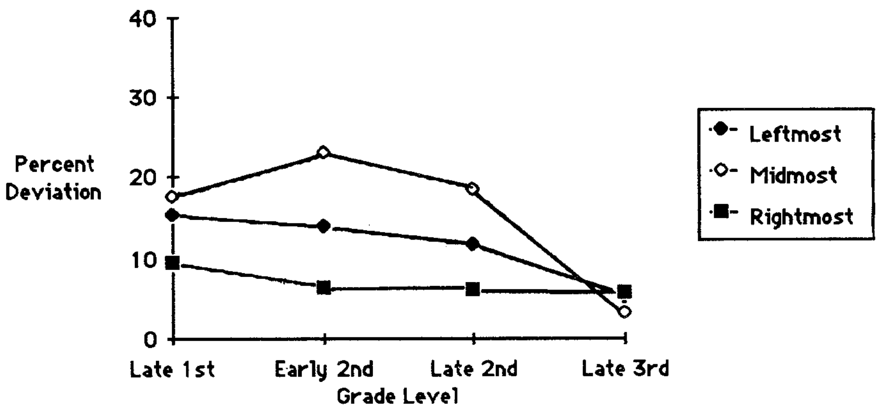

FIGURE 2 Average deviation scores for the 0-to-10 range.

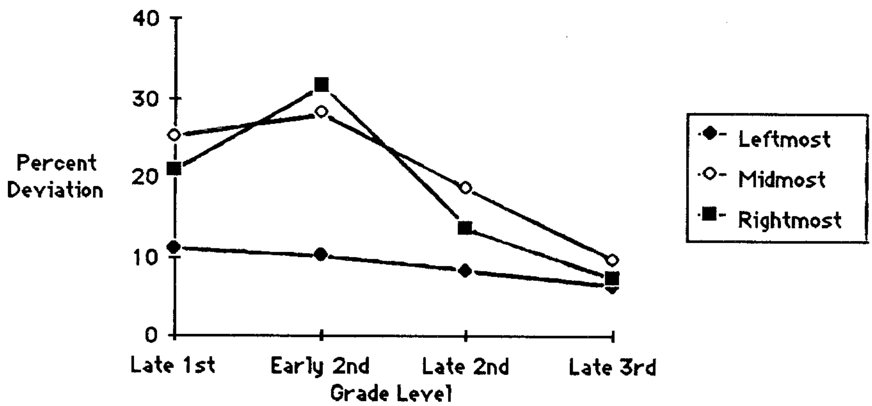

FIGURE 3 Average deviation scores for the 0-to-100 range.

TABLE 4
Contrasts on Deviation Scores Between Grade Levels That Are Significant at $p < .05$ by Scheffe's Test

	Grade Level		
Grade Level	*Late 1st*	*Early 2nd*	*Late 2nd*
Early 2nd	—	—	—
Late 2nd	—	right 100	—
Late 3rd	mid 10 left 100 mid 100	mid 10 left 100 mid 100 right 100	mid 10

of variance (ANOVA) with grouping factors (i.e., Grade and Conservation) and within subject factors (i.e., Range and Target Position) in a 4 (Grade) × 2 (Conservation) × 2 (Range) × 3 (Position) interaction. The deviation scores for 7 of the children[2] could not be used in this, and subsequent ANOVAs, because of extreme values or off-scale solutions on one or more of the targets. The performance of the 7 children omitted here was interpretable in terms of strategies, however, and they were included in the subsequent strategy analyses.

The results of the ANOVA are consistent with the expected sequence-to-proportion shift. First, the significant main effect of grade confirmed that children's deviation scores do decline significantly over the grade levels tested, $F(3, 69) = 10.64$, $p < .0001$. A significant main effect of position, $F(2, 138) = 19.49$, $p < .0001$, and a significant Grade × Position interaction, $F(6, 138) = 2.71$ $p < .02$, showed that the pattern of deviation scores on different target positions changed over time.

The abruptness of change of deviation scores was expected to differentiate between the component skills and conceptual differentiation hypotheses. The analysis of the pattern of change over the four grade levels produced ambiguous results, however. Scheffe tests separately compared each pair of the four grade levels on deviation scores for each of the six numberline problems. These comparisons are summarized in Table 4.

There were no significant differences between late-first and early-second or between late-first and late-second grades, and only one difference, the rightmost target position in the 0-to-100 range, between early- and late-second grades ($p < .05$). In contrast, there were many significant differences between late-third grade and the earliest three grade levels; the midtarget on the 0-to-10 range contrasted with all three of the earlier grade levels, the leftmost target and the mid-targets on the 0-to-100 range contrasted with

[2]There were 3 in late-first, 2 in early-second, 1 in late-second, and 1 in late-third grades, leaving 74 children: 13 in late-first, 21 in early-second, 20 in late-second, and 20 in late-third grades.

late-first and early-second grades; and the rightmost target on the 0-to-100 range contrasted with early-second grade. Note that, with two exceptions, all significant contrasts were between groups separated by at least 12 months of schooling, suggesting a gradual rather than abrupt change over time. This pattern was consistent with the component skills hypothesis but failed to support the conceptual differentiation hypothesis.

Range effects were expected to provide evidence for the component skills hypothesis. Although there was a significant main effect of range, $F(1, 69) = 18.18$, $p < .0001$, with lower deviation scores on the 0-to-10 range than on the 0-to-100 range, the two-way Range × Grade interaction was not significant. Improvements in accuracy over time could not be attributed to greater familiarity with larger numbers in any simple way, and so this result failed to support the component skills hypothesis. However, the three-way Range × Position × Grade interaction was highly significant, $F(6, 138) = 3.74$, $p < .01$, indicating that improvement over time on different numerical ranges was related to target position in complex ways.

In the 0-to-100 range, deviation scores for the mid- and rightmost targets were high in early grades and dropped to a low level by late-third grade. This pattern was consistent with the hypothesized sequence-to-proportion shift. Strategies that rely on unidirectional sequencing would most often be oriented to the 0 end of the line, closest to the leftmost target, and would produce greater errors for the targets farther away from that endpoint. When children began to use proportion relationships, percent deviation scores for the mid- and rightmost targets were reduced and approached the scores for the leftmost targets.

In the 0-to-10 range, a different pattern emerged. The different patterns of results for the two numerical ranges were reflected in a significant Range × Position interaction, $F(2, 138) = 14.80$, $p < .0001$. Closer examination of the deviation scores on the 0-to-10 range showed a strong bias to overestimate the value of the leftmost target (66% overestimated, 4% underestimated) but a much smaller bias to underestimate the value of the rightmost target (18% overestimated, 36% underestimated). These unexpected differential biases undoubtedly contributed to the relatively high deviation scores for that target position but cannot be explained by any of the hypotheses put forward in this article.

The ANOVA showed no significant main effect of conservation and no significant interaction between conservation and any other factor. Hence, there was no direct evidence that conceptual changes in children's conceptualization of linear space contributed to the observed changes in patterns of deviation scores over time.

Strategies. Children's strategies were scored according to four categories: counting from left to right, counting from right to left, using the midpoint position and value, and estimating with no visible strategy used.

TABLE 5
Percentage of Children in Each Strategy Category for 0-to-10 Range Problems

Strategy Category	Grade Level			
	Late 1st	*Early 2nd*	*Late 2nd*	*Late 3rd*
Unidirectional counting				
No counting	18.75	18.18	27.27	23.81
Appropriate	50.00	40.91	45.45	61.90
Inappropriate	31.25	40.91	27.27	14.29
Midpoint strategy				
Use midpoint	12.50	9.09	22.73	47.62

When counting was observed, a second-level score was given based on the counting interval used; counting intervals were usually either 1s or 10s. Three children were observed to count by 5s, and these instances were grouped with those counting by 10s in the following analyses. Counting was always observed to be unidirectional, that is, children counted from one known position value (usually one endpoint but sometimes the midpoint, using the value 5 or 50, depending on numerical range) to a target, but they never used the target's position relative to any other known position to check or to adjust the result of the count. Children sometimes used different strategies on different problems and sometimes used combinations of strategy components on a single problem.

The children's strategy components were scored by the tester at the time of the test, and no videotapes were taken. These same strategy components were identified and scored during a previous pilot testing that was videotaped. The percentage of agreement between two independent raters who scored strategies from the videotapes for 11 pilot children averaged 97%, ranging from 93% to 100%.

Counting was used by a great majority of children at all grade levels. Tables 5 and 6 show that the percentage of children who never used counting on any of the three problems in each numerical range is quite low in all groups. What appears to change over time is the way that counting is used. Two dimensions in the use of counting are of concern here: the direction of counting and the counting intervals. These two dimensions will be examined separately.

Children's counting strategies can be characterized as an estimation in which one starts with a nearby, known position value and counts from there toward the target. Such a strategy is most appropriate when counting begins at a known position close to the target, rather than one considerably farther away. For each numerical range, each child was assigned to either an "appropriate" or "inappropriate" counting category based on the relationship between the direction of counting and the position of the target. No child was assigned to both categories in any one numerical range. A child was

TABLE 6
Percentage of Children in Each Strategy Category for 0-to-100 Range Problems

Strategy Category	Grade Level			
	Late 1st	*Early 2nd*	*Late 2nd*	*Late 3rd*
Unidirectional counting				
No counting	18.75	18.18	27.27	23.81
Appropriate	50.00	36.36	59.09	61.90
Inappropriate	31.25	45.45	13.64	14.29
Midpoint strategy				
Use midpoint	18.75	4.55	22.73	52.38
Interval counting				
Count by 1s	50.00	72.73	36.36	19.05
Count by 10s	25.00	22.73	18.18	61.90
No visible strategy	25.00	4.55	40.91	19.05

placed in the inappropriate category if that child had counted either from the left endpoint to the rightmost target or from the right endpoint to the leftmost target. A child was placed in the appropriate category if he or she was not in the inappropriate category and if he or she had counted from the left endpoint to the leftmost- or midtarget, or from the right endpoint to the rightmost- or midtarget. As defined here, inappropriate use of counting suggests an inflexible unidirectional counting strategy that is insensitive to task circumstances. Tables 5 and 6 show the percentage of children in each of these categories at each grade level. These data indicate that inappropriate counting is relatively frequent at lower grades and drops sharply by late-second grade and that appropriate counting strategies occur at higher frequencies throughout.

For the 0-to-100 range problems, children were scored as using counting by 1s if they counted by 1s on at least two of the three problems in the 0-to-100 range and were scored as using counting by 10s if they counted by 10s on at least two of the three problems in the 0-to-100 range (see Table 6). Children were observed to use only one counting interval on any one problem and were never observed to count by 10s on a 0-to-10 range problem. As shown in Table 6, counting by 1s decreases from high levels at late-first and early-second grades to the lowest level at third grade, whereas counting by 10s is infrequent until late-third grade.

A category termed "no visible strategy" was used to include those children who used no visible strategy on at least two problems in the 0-to-100 range (see Table 6). These children might have used counting or other strategies covertly, or they might have estimated, guessed, or picked an arbitrary favorite number. Thus, this category is not meaningful except to account for the children not included in other categories. Each child could only be scored in one of the three categories in the last three lines of Table 6. However, the percentage of children in each of these three categories occasionally totals less than 100%, because a few children used a different strategy on each of the three problems and, therefore, do not fall into any of the three categories.

One noncounting strategy component, the use of the midpoint value, was also analyzed. When midpoint values were used, they were always used to find the value of the midtarget. Children were scored as either using or not using midpoint values for each of the numerical ranges. The percentage of children who were scored as using the midpoint strategy is shown in Tables 5 and 6. These data show a marked increase in use of the midpoint values across grades, particularly from late-second to late-third grade.

Chi-square tests were used to compare the number of children scored in each of the strategy categories previously described (Tables 5 & 6) across the four grade levels. The analysis of the "no counting" category revealed no significant difference in the extent to which counting was or was not used across all four grade levels for either numerical range. However, the appropriateness of the direction of counting and the counting intervals used did vary across grade levels. On the 0-to-100 range problems, younger children used an inappropriate counting direction significantly more often than did older children, $\chi^2(3, N = 81) = 7.817$, $p < .05$, but there were no significant differences in use of counting in an appropriate direction. Table 5 shows that data for the 0-to-10 range parallel those for the 0-to-100 range, but chi-square analyses did not reach significance. On the 0-to-100 range problems, use of counting by 1s decreased significantly, $\chi^2(3, N = 81) = 14.034$, $p < .003$, and counting by 10s increased significantly over the grade levels tested, $\chi^2(3, N = 81) = 11.771$, $p < .01$. The use of midpoints increased significantly across grades, $\chi^2(3, N = 81) = 13.81$, $p < .003$, on the 0-to-100 range line. Use of the midpoint for 0-to-10 range problems parallels that in the 0-to-100 range, but the data are too sparse for analysis. Thus, though unidirectional counting remained a prevalent strategy across all grade levels, the direction of counting became more flexibly adapted to problem situations in later grades, and unidirectional counting by 1s was supplemented by the increased use of counting by 10s and a new midpoint strategy.

To determine whether these changes were due to incremental acquisition of specific strategies or to conceptual differentiation, the patterns of strategy change over time were assessed. Chi-square analyses were used for pairwise comparisons between adjacent grade levels for all strategies that had shown significant changes across all four grades. These comparisons showed that from early- to late-second grade, there were significant decreases in inappropriate counting, $\chi^2(1, N = 44) = 5.350$, $p < .02$, and in counting by 1s on the 0-to-100 range lines, $\chi^2(1, N = 44) = 5.867$, $p < .02$. From late-second to late-third grades, there were significant increases in counting by 10s, $\chi^2(1, N = 43) = 8.592$, $p < .01$, and in the use of midpoint strategies, $\chi^2(1, N = 43) = 4.040$, $p < .05$.

These results suggest two phases in the shift to more effective strategies: First, there is a drop in the use of inappropriate and ineffective modes of counting, and, somewhat later, there is an increase in the use of new strategy components (i.e., counting by 5s and 10s) and using the value of the midpoint. These results are consistent with the conceptual differentiation view that advances in children's conceptualization of linear space might cause them to reject the use of inappropriate strategies even though they are not yet ready to apply new strategies based on more advanced counting and arithmetic skills. Taken alone, however, these results cannot be unambiguously interpreted.

In order to test directly the conceptual differentiation hypothesis, chi-square analyses of the relationships between conservation and each strategy category in Tables 5 and 6 were conducted. These showed no significant relationship between conservation and strategy use, except for the use of counting by 1s and 10s in the 0-to-100 range; conservers were less likely to use counting by 1s, $\chi^2(1, N = 81) = 3.773, p < .05$, and more likely to use counting by 10s, $\chi^2(1, N = 81) = 3.742, p = .05$. Because both of these strategy variables were also significantly related to grade level, it was necessary to eliminate the possibility that grade and conservation are confounded. Therefore, counting by 1s and 10s was compared for conservers and nonconservers among children in the two second-grade groups only. This analysis showed no significant relationships between counting by 1s or 10s and conservation among second graders. These results failed to support the conceptual differentiation explanation for the strategy shifts reported previously.

In summary, these observations indicate that there are qualitative changes in children's strategies over the four grade levels tested, and these changes indicate a shift from simple unidirectional counting by 1s in the earliest grades to the use of midpoint values and alternative counting intervals by the end of third grade. These shifts appear to be related to incremental acquisition of component skills. Little evidence was found to support a conceptual differentiation hypothesis.

It is striking that the use of counting by 10s does not increase significantly until over a year after children master the verbal counting skill. This finding might support a conceptual differentiation argument, or it might mean that children's counting by 10s is initially a verbal string with little or no concrete significance. Children have a great deal of experience counting objects by 1s and relatively little experience counting groups of objects by 10s. Therefore, the delay in use of counting by 10s could be caused by the younger children's lack of understanding of the concrete significance of the verbal string.

EXPERIMENT 2

Experiment 2 was designed to examine a more sophisticated concept of subdivision: the subdivision of a line into equal-size portions. Piagetian theory predicts that such a concept would be another important aspect of the child's developing understanding of measurement. This concept is here termed the *equal interval principle.*

> True measurement of distance and lengths begins when the subject recognizes that any length may be decomposed into a series of intervals which are known to be equal because one of them may be applied to each of the others in turn. (Piaget et al., 1960, p. 149)

The equal interval principle holds that the distance between any two successive numbers must be the same as the distance between any other pair of successive numbers. In this experiment, a distinction is made between principled and conventional concepts of equal interval units. A conventional concept of equal intervals is obtained when an individual is able to distinguish an equal interval ruler from others and to recognize that rulers are usually made that way. The individual is said to have a principled concept only when he or she further recognizes the functional significance of equal intervals. That is, rulers really must be marked off in equal intervals for a systematic measurement system to work, and other configurations are not acceptable.

As shown in Table 1, in first grade, children in this sample learned to construct rulers and to assess length by repeated application of a single standard unit. This use of repeated units is presumably intended to introduce the concept of a whole measure constructed of equal units. Children might not view such constructions as a unified "whole" in the same sense as a single undivided object, however, but only as the trace or record of a procedure for counting "steps." In third grade, however, children are introduced to the concept of fractions in terms of subdivision of wholes into equal parts. This lesson is more directly related to subdivision and, in particular, equal interval subdivisions than is the construction of a whole from repeated application of a single unit. Thus, the third-grade curriculum might directly affect children's understanding of equal interval subdivision.

According to Piagetian theory, nonconservers might not infer that a standard unit marks off the same interval each time it is moved to a new position. Thus, nonconservers might not pick up an equal interval concept directly from this mode of instruction. If this is correct, conservation of length might be a necessary prerequisite for the recognition of equal intervals on a ruler.

The sequence of lessons shown in Table 1 and Piagetian theory suggest that two major elements that could lead to a principled concept of equal interval units are present by the spring of third grade: Explicit lessons are presented in third grade on subdivision of wholes into equal-size portions, and most third graders conserve length (see Table 2).

Children who recognize equal intervals as a conventional but not a strong constraint might abandon the equal interval configuration for another configuration under conditions that make the equal interval configuration inconvenient. Thus, there are three relevant aspects of equal interval rulers to be examined here: the ability to discriminate an equal interval configuration from other configurations, the acknowledgment that equal interval rulers are usually the "right" ones to use, and the persistent use of equal interval rulers under conditions that make them inconvenient.

Method

Subjects and procedure. Each of the children in Experiment 1 was tested approximately 1 week later for Experiment 2. Children were tested individually by the same person as in Experiment 1 and were given a two-part ruler discrimination test followed by a ruler preference test. They were then asked to use the rulers to solve the same estimation problems that were presented in Experiment 1.

The ruler discrimination test was done first with five 0-to-10 range rulers and then with five 0-to-100 range rulers. Within each range, the five rulers varied the configuration of interval sizes between numbers. Figure 4 shows the five-unit configurations with the 0-to-100 range rulers. The purposes of this test were to ensure that each child noticed the interval configurations and to assess the children's ability to discriminate among the rulers on the basis of those configurations. For each set, the five rulers were placed in front of the children, aligned one above the other (as in Figure 4) but randomly ordered. They were dumped out of an envelope onto a table for each subject, and no attempt was made to record the position of each configuration.

The instructions for the ruler discrimination test were:

> Here are five rulers. Each one is different. I am going to pick one in my mind, but I'm not going to show you which one. From the words I use to describe it, you tell me which one I am thinking of.

The following descriptions were then presented:

> The ruler I am thinking of has numbers that start out close together then get farther and farther apart; The ruler I am thinking of has numbers spaced far apart in the beginning but they get closer and closer together; In the ruler I am thinking of, the spaces between the numbers are all the same size; In the ruler I am thinking of, the numbers are spaced close, far, close, far all the way along the ruler.

The irregular interval ruler was never described, although it was always included in the five rulers displayed. Each child was scored for the number correct out of the eight discrimination questions (four for each numerical range).

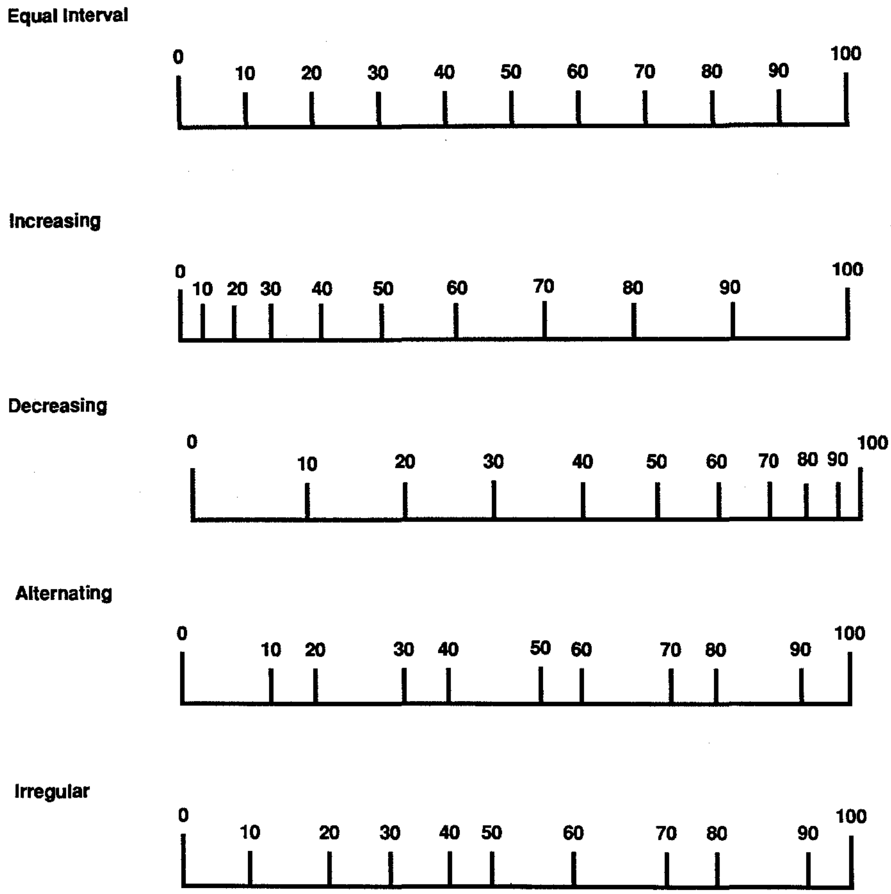

FIGURE 4 The five ruled-line configurations.

In the ruler preference test, all 10 rulers were placed in front of the child, and the child was asked: "Remember the numberline problems you did last week?" [Children always said "yes."] "Do you think these rulers would help you solve those numberline problems?" [Children always agreed that they would help.] The children were shown the first numberline problem and were then asked, "If you were going to use these rulers to help you solve the numberline problems, which one would be best to use?

Immediately after this question, six blank numberline problems similar to those used in Experiment 1 were presented to each child individually and in the same order as in Experiment 1. Before the entire set of six, the child was instructed: "Use these rulers to help you solve these numberline problems." After each problem was solved, the tester made sure that the child returned the ruler he or she had used to the array of rulers on the table. The target po-

TABLE 7
Percentage of Children Who Chose the Equal Interval Ruler as Best and Who Used the Equal Interval Ruler Consistently

	Grade Level			
Children's Choice	*Late 1st*	*Early 2nd*	*Late 2nd*	*Late 3rd*
Chose as best	37.5	68.2	54.6	76.2
Used consistently	12.5	18.2	13.6	42.9

sitions used in Experiment 2 for the 0-to-10 range problems were 1.6, 4.4, and 7.8; those for the 0-to-100 range were the same as in Experiment 1.

None of the targets exactly matched the position of any ruler marking on any of the ruler configurations when the 0 endpoint of the ruler and numberline were properly aligned. This was done so that no matter which ruler a child initially tried, he or she would have to resolve the mismatch between the marked positions and the target position in some way. It was expected that this mismatch situation would be sufficient to cause children who did not hold a strong principle for the use of equal intervals to shift to other rulers in search of ones that might better match. Because no ruler ever precisely matched any target position, this could not result in a bias for any particular ruler configuration, and so any bias observed in consistent use of rulers could be attributed to the child's bias for that configuration based on underlying assumptions about ruler configurations in general.

Results and Discussion

Ruler discrimination. Out of the eight discrimination questions, children averaged 6.2 correct answers ($SD = 1.5$). A two-way ANOVA comparing number correct over grade and conservation level showed no significant main effects or interactions.

Ruler preference. Ruler preference questions were intended to assess knowledge of the equal interval principle. In response to the question, "Which ruler would be best," children either selected one of the equal interval rulers or claimed that no one ruler was better than any other. No child ever selected a ruler with any other unit configuration as "best." Although there was an increase with grade in the frequency of selecting the equal interval ruler as best on this task (see Table 7), a chi-square analysis showed only a nonsignificant trend for differences between grades, $\chi^2(3) = 6.58, p < .09$. Chi-square also showed no significant relationship between conservation and selecting the equal interval ruler as best. These results indicate that neither the fraction lessons in third grade nor conservation of length is a necessary prerequisite for recognizing equal interval subdivisions as a relevant convention.

TABLE 8
Number of Children Using the Equal Interval Ruler Consistently Who Conserved and/or Chose the Equal Interval Ruler as Best or Did Neither

Children's Choice	*Consistent Use*	*No Consistent Use*
Neither	0	16
Choose, not conserve	2	13
Conserve, not choose	2	14
Both	14	20

Use of rulers for estimating target values. When using the rulers to solve the numberline estimation problems, all children showed familiarity with the use of rulers for measurement. They consistently lined up the "zero mark" of the ruler with the left endpoint of the numberline, then looked for a marked ruler position at or close to the target on the numberline.

The use of equal interval or other rulers was intended to identify those children with a principled concept of equal interval units. Children with a principled concept of equal intervals were expected to select and to use the equal interval ruler on all six of the numberline problems. Other children were expected to select ruler configurations at random.

Analysis of the frequency of each child's selection and use of each ruler type showed that 18 children always used the equal interval ruler. The Kolmogorov–Smirnov one-sample test (Siegel, 1956) showed that, among the 63 children who did not always select the equal interval ruler, the frequency distribution of ruler use was not different from chance for any ruler configuration; however, among the total population of 81 children, the frequency of consistent use of the equal interval ruler was significantly nonrandom ($N = 78$, $D = .299$, $p < .01$). Based on this analysis, children were identified as either consistent (six of six) or inconsistent (less than six) users of equal interval rulers. The status of two children who used the equal interval rulers five of the six times is arguable, but the more conservative alternative is to identify them with the inconsistent users.

Subsequent analyses showed that, in contrast to conventional recognition of equal interval units, consistent use—suggesting principled understanding—is related to grade level, conservation of length, and the achievement of conventional recognition. Table 7 shows the percentage of children at each grade who were consistent users of equal interval rulers in the ruler-use task. The frequency of consistent users was too small to permit a chi-square comparison over all four grades together; therefore, comparisons of the number of consistent equal interval users were made across each pair of adjacent grades. These comparisons showed a significant increase from late-second to late-third grade, $\chi^2(1, N = 81) = 4.56$, $p < .04$.

Table 8 summarizes the relationships among consistent use of the equal interval ruler, the four possible combinations of achievements in conservation of length, and choosing the equal interval ruler as best. Although choosing the equal interval ruler as best and conservation of length were found to be unrelated to each other, each appeared to be related to the consistent use of the equal interval ruler. A chi-square analysis on the frequencies in Table 8 (collapsing the category "neither" with "choose, not conserve" and collapsing "conserve, not choose" with "both") showed that conservation of length was significantly related to consistent use of equal interval rulers, $\chi^2(1, N = 81) = 7.23, p < .01$; specifically, virtually no one who was a consistent user did not conserve, but conserving did not guarantee consistent use. In order to ensure that this effect was not the result of confounding grade and conservation levels, the same chi-square was run on second graders only, pooling early- and late-second graders together. This analysis showed that the same relationship held among second graders, $\chi^2(1, N = 44) = 3.73, p < .05$.

Another chi-square on the frequencies in Table 8 (collapsing the category "neither" with "conserve, not choose" and collapsing "both" with "choose, not conserve") showed that choosing the equal interval ruler as best was itself significantly related to its consistent use, $\chi^2(1, N = 81) = 7.81, p < .01$. A conventional recognition of equal intervals appears to precede its consistent application.

According to these results, both conventional recognition of equal intervals and conservation of length are important for the acquisition of the equal interval principle as assessed by consistent use of the equal interval ruler. This suggests that there might be an additive relationship between these two precursors; children must both notice that conventional rulers have equal intervals and achieve conservation of length before acquiring a principle that sustains consistent use.

To test this possibility, a chi-square was done on the frequencies in Table 8 collapsing the categories "choose, not conserve" and "conserve, not choose," leaving three categories as follows: (a) neither conserved length nor showed preference for any ruler, (b) either conserved or preferred the equal interval ruler but not both, and (c) both conserved and preferred the equal interval ruler. This chi-square showed highly significant relationships, $\chi^2(2) = 13.20, p < .001$; 14 of the 18 consistent users were among those who both conserved length and preferred equal interval rulers. Note, however, that more than half of the children in the "both" category still did not use the equal interval ruler consistently. Conservation of length and conventional recognition might be precursors of the equal interval principle, but they are not sufficient to ensure it.

CONCLUSION

Over Grades 1 through 3, children appeared to approach the numberline estimation tasks in Experiment 1 at first in a purely sequential way and later with increasing attention to spatial configurations related to numerical quantity on the numberline; thus, children shifted from sequential to proportional reasoning over the first three elementary school grades. Evidence from accuracy and strategy observations indicated that this sequence-to-proportion shift was accomplished through the acquisition of component arithmetic and counting skills, as well as the increasing ability to judiciously apply those skills to varying problem situations. The evidence showed no clear need for conceptual differentiations in children's understanding of linear space for the development of at least a rudimentary sense of proportion. There was evidence, however, that the achievement of conservation of length is one important component in the acquisition of a principled concept of equal intervals.

Although children who do not conserve length are quite capable of perceptually discriminating an equal interval configuration from other configurations, a recognition of the functional significance of the equal interval configuration in a measurement system appears to depend upon several prior accomplishments. The results of Experiment 2 point to two of them: the knowledge of an equal interval convention, and the differentiated concepts of position and length as assessed by the conservation task. That there are others as yet undiscovered is demonstrated by the sizable proportion of children who were shown to have both of these achievements but who did not demonstrate the equal interval principle.

With this it is possible to begin to interpret the early observation of children's inappropriately tiny adjustments in a task involving successive approximations on a numberline. Young children at first may have no real conceptualization of proportion or scale in a numberline representation and have limited ability to apply what verbal counting and arithmetic skills they might have. Without an adequate understanding that the numberline is divided into equal-size units, there is no way for the child to know that each increment will be the same size as the one before it. The sense of proportion that adults experience with respect to numerical values is a gradual and complex development requiring a variety of educational experiences and cognitive achievements.

ACKNOWLEDGMENTS

I gratefully acknowledge Louisa Pierson for her help in data collection and processing and Craig Barclay, Sharon Carver, Lucia French, Lauren Resnick, and two anonymous reviewers for their many helpful suggestions on drafts of this article.

REFERENCES

Ashcraft, M. H., & Battaglia, J. (1978). Cognitive arithmetic: Evidence for retrieval and decision processes in mental addition. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 527–538.

Ashcraft, M. H., & Fierman, B. A. (1982). Mental addition in third, fourth and sixth graders. *Journal of Experimental Child Psychology, 33*, 216–234.

Ashcraft, M. H., & Stazyk, E. H. (1981). Mental addition: A test of three verification models. *Memory & Cognition, 9*, 185–196.

Fuson, K. (1984). More complexities in subtraction. *Journal for Research in Mathematics Education, 15*, 214–225.

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

Levin, J. A., Boruta, M. J., & Petitto, A. L. (1983). How do children think about numbers? Let us count the ways. In P. J. Hayes (Ed.), *Proceedings of the Fifth Annual Conference of the Cognitive Science Society*. Rochester, NY: University of Rochester.

Miller, K. F., & Gelman, R. (1983). The child's representation of number: A multidimensional scaling analysis. *Child Development, 54*, 1470–1779.

Murray, P. L., & Mayer, R. E. (1988). Preschool children's judgements of number magnitude. *Journal of Educational Psychology, 80*, 206–209.

New York State Education Department. (1985). *Mathematics K–6: A recommended program*. Albany, NY: The State Education Department.

Piaget, J., Inhelder, B., & Szeminska, A. (1960). *The child's conception of geometry*. New York: Basic.

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.