Research Article

THE DEVELOPMENT OF NUMERICAL ESTIMATION: Evidence for Multiple Representations of Numerical Quantity

Robert S. Siegler and John E. Opfer
Carnegie Mellon University

Abstract—*We examined children's and adults' numerical estimation and the representations that gave rise to their estimates. The results were inconsistent with two prominent models of numerical representation: the logarithmic-ruler model, which proposes that people of all ages possess a single, logarithmically spaced representation of numbers, and the accumulator model, which proposes that people of all ages represent numbers as linearly increasing magnitudes with scalar variability. Instead, the data indicated that individual children possess multiple numerical representations; that with increasing age and numerical experience, they rely on appropriate representations increasingly often; and that the numerical context influences their choice of representation. The results, obtained with second graders, fourth graders, sixth graders, and adults who performed two estimation tasks in two numerical contexts, strongly suggest that one cause of children's difficulties with estimation is reliance on logarithmic representations of numerical magnitudes in situations in which accurate estimation requires reliance on linear representations.*

Estimation is a pervasive process in both school and everyday life. It is also a process that many children find difficult. This difficulty has been ascribed to inadequate central conceptual structures, mindless symbol manipulation, lack of number sense, and misunderstanding of arithmetic (Hiebert & Wearne, 1986; Joram, Subrahmanyam, & Gelman, 1998; Sowder & Wheeler, 1989).

The present research suggests a different, and rather surprising, source of children's estimation difficulties: flawed understanding of the decimal system, leading to use of inappropriate numerical representations. Children appear to possess multiple representations of numerical magnitude and to rely fairly often on early-developing logarithmic representations in situations in which accurate estimation requires reliance on later-acquired linear representations.

Research on how representations of magnitude influence estimation began with Fechner's Law, which states that the magnitude of a sensation is a logarithmic function of objective stimulus intensity. A number of investigators have suggested that this law describes representations of numerical as well as physical magnitudes (Banks & Hill, 1974; Dehaene, 1997). Consistent with the implication of Fechner's Law, adults' speed and preschoolers' accuracy and speed in comparing numerical magnitudes decrease logarithmically as the ratio of the numbers approaches 1 (Dehaene, Dupoux, & Mehler, 1990; Sekuler & Mierkiewicz, 1977). Similar *distance effects* are shown by infants in habituation experiments and by a variety of nonhuman animals in conditioning paradigms designed to parallel numerical estimation (Washburn & Rumbaugh, 1991; Xu & Spelke, 2000). Adults' speed and preschoolers' accuracy also decrease logarithmically with numerical magnitude when they are comparing numbers separated by equal distances (Moyer & Landauer, 1967; Sekuler & Mierkiewicz, 1977). Again, similar *size effects* have been observed with infants and nonhuman animals (Dehaene, Dehaene-Lambertz, & Cohen, 1998; Starkey & Cooper, 1980).

To explain these data, Dehaene (1997) proposed the logarithmic-ruler model:

> Each time we are confronted with an Arabic numeral, our brain cannot but treat it as an analogical quantity and represent it mentally with decreasing precision, pretty much as a rat or chimpanzee would do . . . . Our brain represents quantities in a fashion not unlike the logarithmic scale on a slide rule, where equal space is allocated to the interval between 1 and 2, between 2 and 4, and between 4 and 8. (pp. 73, 76)

Within Dehaene's model, the logarithmic data patterns in previous experiments reflect the underlying representation of numbers. Reliance on these representations "occurs as a reflex" (p. 78) and cannot be inhibited.

Gibbon and Church (1981) proposed a different account of numerical representation: the accumulator model. They suggested that people and other animals represent quantities, including numbers, as equally spaced, linearly increasing magnitudes with scalar variability. Gallistel and Gelman (2000) explained scalar variability as follows:

> The non-verbal representatives of number are mental magnitudes (real numbers) with scalar variability. Scalar variability means that the signals encoding these magnitudes are 'noisy'; they vary from trial to trial, with the width of the signal distribution increasing in proportion to its mean. (p. 59)

Within the accumulator model, the logarithmic data patterns reflect degree of overlap between representations. Representations of number entail higher scalar variability with increasing magnitude; therefore, comparisons at any given numerical distance will be slower and less accurate the larger the magnitude. Similarly, representations of magnitudes that are closer in size will overlap more and therefore be harder to discriminate than representations of magnitudes that are farther apart in size. Consistent with this analysis, the coefficient of variation (*SD*/number being estimated) of numerical estimates has been found to be constant over numerical size for both rats and humans on several estimation tasks (Whalen, Gallistel, & Gelman, 1999). Relying on these data, advocates of the accumulator model have explicitly rejected the hypothesis that people might have other numerical representations. This position is reflected in Brannon, Wusthoff, Gallistel, and Gibbon's (2001) statement, "Given the current state of knowledge, we view the idea that number is represented both linearly and logarithmically as unparsimonious" (p. 243).

Both the logarithmic and the accumulator models provide attractive accounts of how people and other animals represent quantities. However, it seems likely that neither model describes the only representation that people use. Instead, we believe that (a) individual children know and use multiple representations of numerical quantity over

Address correspondence to Robert S. Siegler or John E. Opfer, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213; e-mail: rs7k@andrew.cmu.edu or jopfer@andrew.cmu.edu.

a period of many years; (b) with development, children rely on formally appropriate representations in an increasing range of numerical contexts; and (c) the numerical context influences children's choice of representation. The main purpose of the study we report in this article was to test these hypotheses.

The perspective on development of numerical representations suggested by these three hypotheses differs considerably from the perspective suggested by prior accounts. Both the logarithmic-ruler and the accumulator models depict people and other animals as invariably representing numerical magnitudes in a single way. In many other domains, however, multiple strategies and representations have been found to coexist over prolonged periods of development (Goldin-Meadow, 2001; Keil, 1989; Siegler, 1996). The same may be true of representations of numerical magnitudes. In addition to the representations implied by the logarithmic-ruler and accumulator models, people may possess a *linear-ruler* representation—a linear representation of quantitative magnitude that does not entail scalar variability. Such linear representations may be formed during middle childhood as a result of experience with numbers and mathematics instruction and may be extended to wider ranges of numbers as children gain experience with them. Thereafter, they may coexist and compete with the logarithmic-ruler and accumulator representations.

Within this overlapping-waves perspective (Siegler, 1996), possessing multiple representations is useful because different representations are optimal in different situations. Thus, logarithmic representations and decelerating power functions are ideal for representing situations in which differences at the low end of the range are paramount (e.g., estimating availability of food), whereas linear representations are ideal when all parts of the range are equally important (e.g., estimating answers to arithmetic problems).

In the current study, we presented second, fourth, and sixth graders and adults with two complementary estimation tasks. On the number-to-position (NP) task, participants were shown a number and asked to estimate its position on a number line. On the position-to-number (PN) task, participants were shown a position on a number line and asked to estimate the number that corresponded to it. Some number lines had 0 at one end and 100 at the other; others had 0 at one end and 1,000 at the other.

These number-line tasks were of interest for several reasons. They embodied the core property of estimation—translation between representations (numerical and spatial)—without demanding specific knowledge of measurement units that might be lacking in young children. They had high ecological validity; children are familiar with number lines because they appear on walls and in textbooks in many classrooms. They also could be used to examine a much wider range of numbers than estimation tasks used in previous research.

Especially important, the number-line tasks provided a test of the three hypothesized representations, because they would yield distinct patterns of means and variances on the tasks. Consistent reliance on a representation with scalar variability implies that both means and variability of estimates should increase linearly with increasing numerical magnitude on both the NP and the PN tasks. Consistent reliance on a logarithmic-ruler representation implies that mean estimates should increase logarithmically with numerical magnitude on the NP task and exponentially with numerical magnitude on the PN task (Fig. 1). Depending on the interpretation of the logarithmic-ruler model, variability on both tasks would either increase linearly with numerical magnitude (Dehaene et al., 1998) or remain constant (Dehaene, 2001). Finally, reliance on a linear-ruler representation would yield linearly increasing means but not scalar variability.

Obtaining estimates for both 0–100 and 0–1,000 number lines and for both the NP and the PN tasks from participants of a wide range of ages allowed us to test our main hypotheses. With regard to the first hypothesis, if participants possessed multiple representations of numerical quantity, they would generate multiple estimation patterns. With regard to the second hypothesis, we expected that age and experience would bring increasing reliance on the linear representation. With regard to the third hypothesis, we expected that the 0–100 scale would elicit linear patterns of estimates more often than the 0–1,000 context, whereas the 0–1,000 context more often would elicit logarithmic patterns. Data consistent with these hypotheses would demonstrate the value of recognizing that children know and use multiple representations of numerical quantity.

METHOD

Participants

There were 32 participants in each of four groups: second graders (mean age = 7.9, *SD* = 0.4), fourth graders (mean age = 9.6, *SD* = 0.3), sixth graders (mean age = 11.8, *SD* = 0.4), and undergraduates. The children attended a suburban school in an upper-middle-class area; the undergraduates were paid volunteers. Two female research assistants served as experimenters.

Tasks

Each problem involved a 25-cm line, with the left end labeled "0" and the right end labeled "100" or "1,000." On the NP task, the number to be estimated appeared 2 cm above the center of the line. On the PN task, the position to be estimated was indicated by a vertical hatch mark that intersected the number line. Two sets of numbers-positions with similar distributions of numbers were created for each scale. For the 0–1,000 scale, Set A included the values 4, 6, 18, 71, 230, and 780; Set B included 2, 6, 25, 86, 390, and 810. For the 0–100 scale, Set A included 2, 4, 6, 18, 42, and 71; Set B included 2, 3, 6, 25, 67, and 86. These numbers were chosen to maximize discriminability of logarithmic and linear functions and to minimize the influence of specific knowledge, such as that 50 is halfway between 0 and 100.

Procedure

Participants were tested in two sessions 1 to 2 days apart. Type of task (NP or PN), set (A or B), and time limit for estimates (4 s or 30 s) were counterbalanced so that each of the eight possible combinations was presented to 4 participants of each age in Session 1. The purpose of varying the time limit was to determine if enabling young children to use time-consuming strategies, such as counting or repeated division, improved their estimates. In Session 2, the statuses of all three variables were reversed; thus, children who were given the PN task with Set A numbers and the 4-s time limit in Session 1 were given the NP task with Set B numbers and the 30-s limit in Session 2. Problems were blocked so that within a session, children answered all the problems on the 0–100 scale before any of the problems on the 0–1,000 scale, or vice versa. Order of the scales was counterbalanced; items within each scale were randomly ordered, separately for each child, and presented in small workbooks, three problems per page.

For the NP task, the experimenter began by saying, "Today we're going to play a game with number lines. What I'm going to ask you to

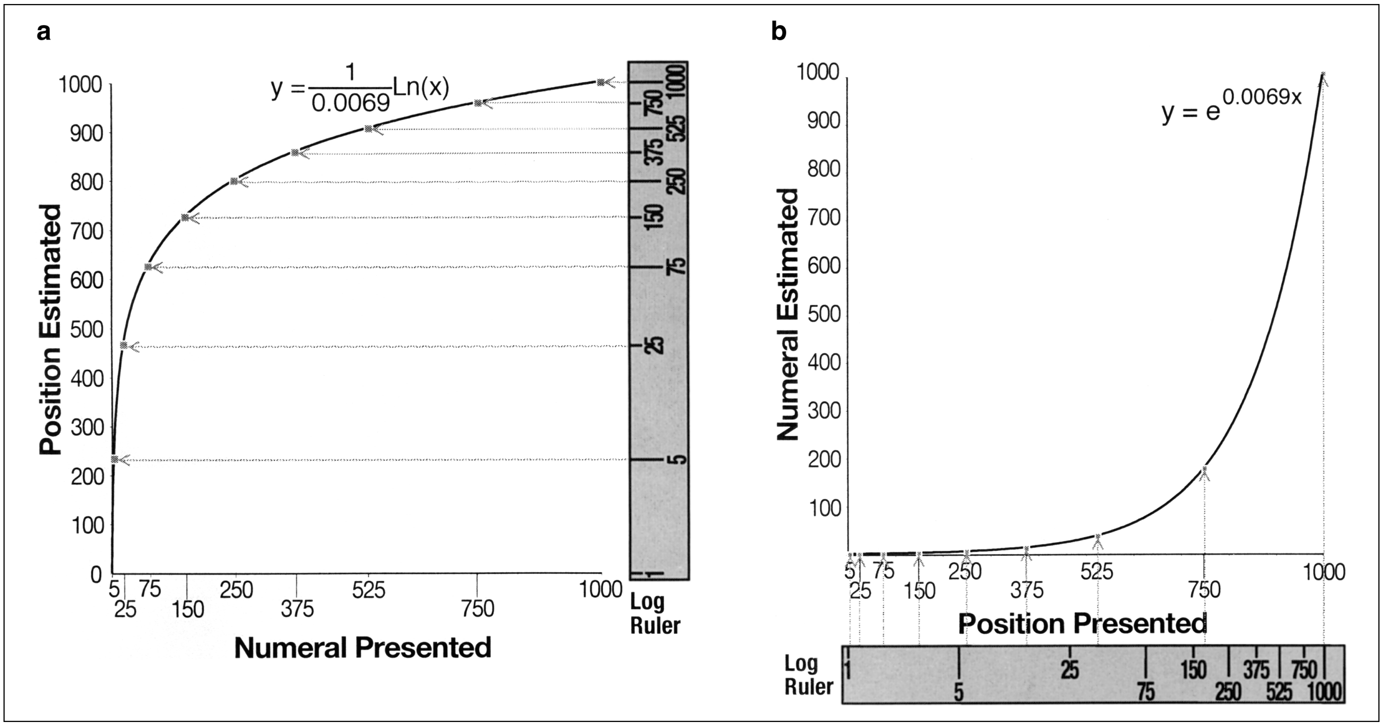

Fig. 1. Predicted estimates of the logarithmic-ruler model for the number-to-position (NP) task and for the position-to-number (PN) task. On the NP task (a), a logarithmic representation of numerical magnitude yields logarithmically increasing estimates of the position that corresponds to each number. For example, a child presented the numeral "75" would estimate the position to be slightly more than 60% of the way from 0 to 1,000. On the PN task (b), a logarithmic representation of numerical magnitude yields exponentially increasing estimates of the number that corresponds to each position. For example, a child presented a position slightly more than one third of the way between 0 and 1,000 would estimate the corresponding numeral to be "20." In both cases, the equation used to generate the function is indicated near the top right of the curve.

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 100 [1,000] at this end. If this is 0 and this is 100 [1,000], where would you put *N*?" (with *N* being the number specified on the particular trial). Instructions for the PN task were similar, except that the experimenter asked before each item, "If this is 0 and this is 100 [1,000], what is this number [pointing to the hatch mark]?"

RESULTS AND DISCUSSION

The primary analyses involved comparisons of the fit of linear, logarithmic, and exponential models to the median estimates for the numerical values. (We also tested a power function model, but it was never the best-fitting equation in any analysis.) Sixteen analyses were conducted, reflecting the possible combinations of age group (second graders, fourth graders, sixth graders, and adults), task (NP or PN), and scale (0–100 or 0–1,000). Preliminary analyses indicated that neither problem set nor time limit interacted with the variables of primary interest (age, task, and scale), and that neither had more than a minimal influence on accuracy; therefore, neither variable is considered further.

The analyses revealed that with age, children's estimates changed substantially. This was especially true on the 0–1,000 number lines. As shown in Figure 2, second graders' median estimates fit the logarithmic model far better than the linear one. On the NP 1,000 (NP task, 0–1,000 number line), the logarithmic equation accounted for 95% of the variance in median estimates, whereas the best-fitting linear equation accounted for 63%, $t(10) = 3.25$, $p < .01$. Complementarily, on the PN 1,000 task (PN task, 0–1,000 number line), the best-fitting exponential equation accounted for 95% of the variance in the median estimates, whereas the best-fitting linear equation accounted for 62%, $t(10) = 2.53$, $p < .05$.

Fourth graders' median estimates were fit equally well by the two models (Fig. 2). On the NP 1,000 task, the logarithmic equation accounted for 93% of the variance in the median estimates, and the linear equation accounted for 82%, $t(10) = 1.28$, n.s. On the PN 1,000 task, the exponential equation accounted for 87% of the variance in fourth graders' median estimates, and the linear equation accounted for 90%, $t(10) = 0.23$, n.s. In contrast, as shown at the bottom of Figure 2, sixth graders' and adults' median estimates were fit better by linear than by logarithmic or exponential functions. On the NP 1,000 task, the logarithmic equation accounted for 78% of the variance in sixth graders' median estimates and 73% of the variance in adults' estimates, whereas the linear equation accounted for 100% of the variance for both age groups: sixth graders, $t(10) = 2.75$, $p = .01$; adults, $t(10) = 4.03$, $p < .005$. On the PN 1,000 task, the exponential equation accounted for 73% of the variance in sixth graders' median estimates and for 66% of the variance in adults' estimates, whereas the linear equation accounted for 97% of the variance for sixth graders

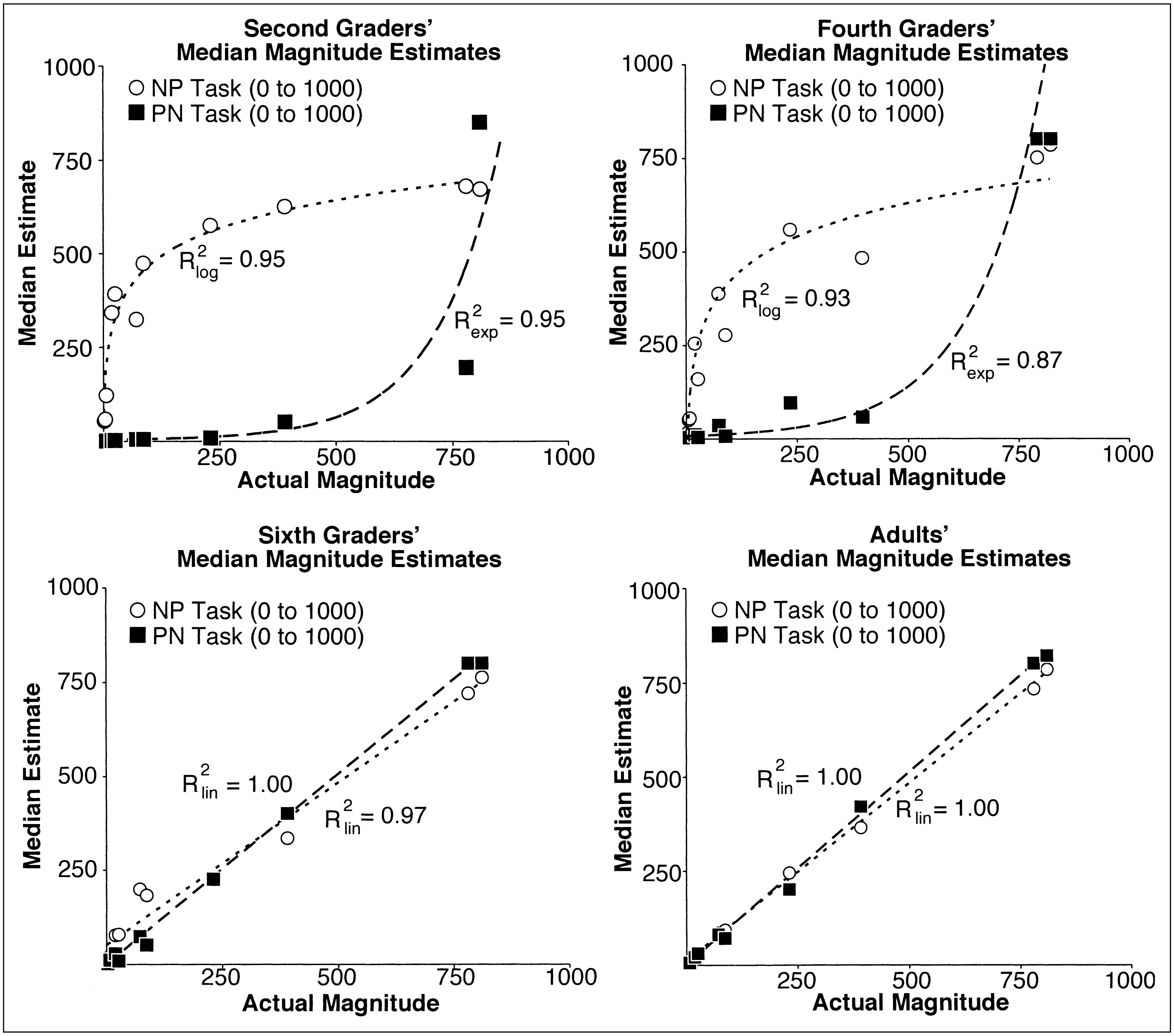

Fig. 2. Median estimates on the number-to-position (NP) and position-to-number (PN) tasks for the 0–1,000 number line. Results are displayed separately for each age group. The dotted and dashed lines indicate the best-fitting functions, and for each function the percentage of variance accounted for (R^2) is indicated.

and 100% of the variance for adults: sixth graders, $t(10) = 1.60$, $p < .10$; adults, $t(10) = 1.61$, $p < .10$.

The finding that second graders' median estimates on the 0–1,000 number line followed a logarithmic pattern is consistent with the logarithmic-ruler model but inconsistent with the accumulator model. The finding that sixth graders' and adults' median estimates followed a linear pattern is consistent with the accumulator model but inconsistent with the logarithmic-ruler model.

To test whether the group-level findings reflected individual performance, we regressed individuals' estimates on each task against objective magnitudes, using logarithmic, exponential, and linear regression formulas. (Because the exponential regression line formula, $y = e^{cx}$, could not be calculated for estimates of zero, we added 1 to both actual and estimated magnitudes.) We assigned a 1 to the model that best fit each participant's estimates and a 0 to the remaining models. Then we conducted a chi-square test to examine the association between age group and the proportion of children for whom each model provided the best fit on each task. (We also tested power functions, but they yielded the best fit for fewer than 15% of second and fourth graders and almost no older individuals.) In the tests of the frequency of the exponential function providing the best fit on the NP 1,000 task and of the frequency of the logarithmic function providing the best fit on the PN

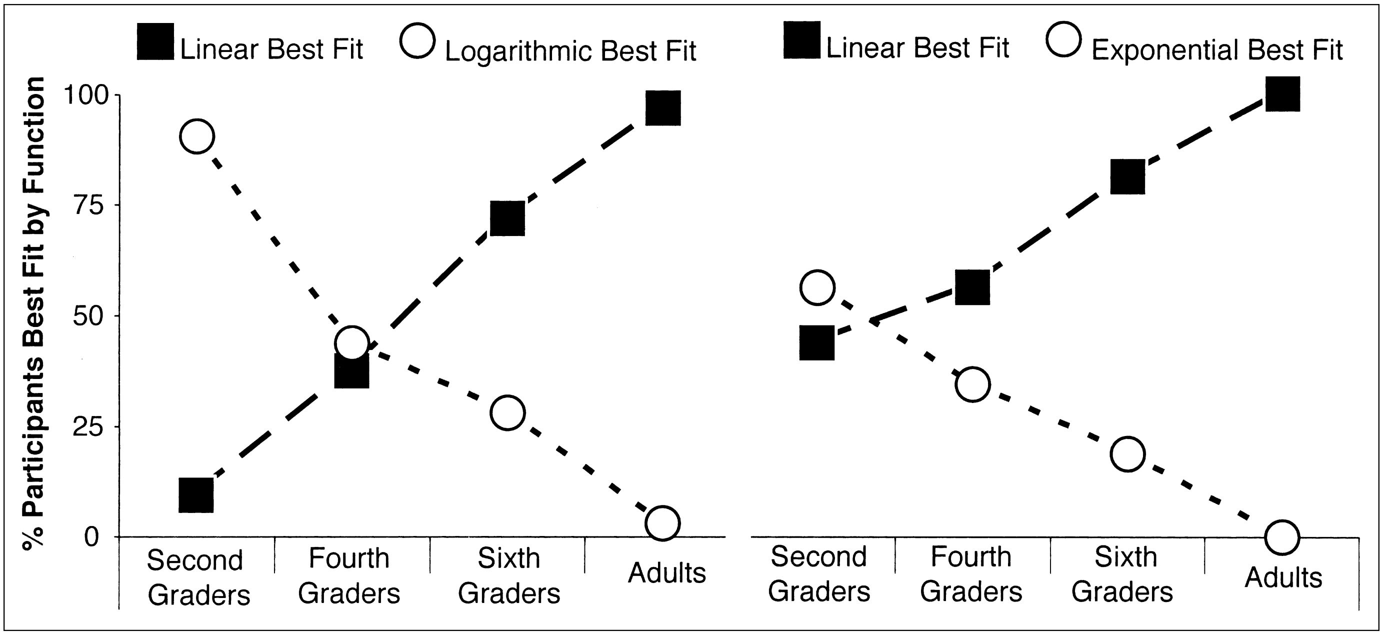

Fig. 3. Percentage of participants in each age group whose estimates on the number-to-position task (left side) and position-to-number task (right side) in the 0–1,000 context were best fit by linear, logarithmic, and exponential patterns.

1,000 task, the expected values were too low to calculate a Pearson chi-square. Therefore, we restricted the analyses to logarithmic and linear models on the NP 1,000 task and to exponential and linear models on the PN 1,000 task.

On the NP 1,000 task, there were significant associations between age and the proportion of children whose estimates best fit the linear and logarithmic models, χ^2s(3, N = 128) = 56.94 and 53.68, respectively, *p*s < .0001. The linear model best fit the estimates of 97% of adults, 72% of sixth graders, 38% of fourth graders, and 9% of second graders (Fig. 3). The logarithmic model best fit the estimates of 91% of second graders, 44% of fourth graders, 28% of sixth graders, and 3% of adults.

On the PN 1,000 task, there was a similar association between age and the best-fitting model, χ^2s(3, N = 128) = 29.19 and 27.49 for linear and exponential models, *p*s < .0001. The linear model best fit the estimates of 100% of adults, 81% of sixth graders, 56% of fourth graders, and 44% of second graders. The exponential model best fit the estimates of 56% of second graders, 34% of fourth graders, 19% of sixth graders, and 0% of adults (Fig. 3). Thus, the data for individual participants paralleled the group data.

Could the results have stemmed from the younger children not understanding the number-line task? If so, then second and fourth graders' estimates on the 0–100 and 0–1,000 lines should be equally nonlinear. In fact, the linear model better fit second and fourth graders' median estimates on the 0–100 number lines than on the 0–1,000 lines, particularly on the NP task: second graders, R^2 = .86, *SD* = .16, vs. R^2= .63, *SD* = .19, $t(31) = 6.67$, $p < .001$; fourth graders, R^2 = .88, *SD* = .21, vs. R^2= .75, *SD* = .20, $t(31) = 3.20$, $p < .01$. Again, individual participants' performance followed the same pattern. The best-fitting equation was linear for more second and fourth graders on the 0–100 line than on the 0–1,000 line (second graders, 56% vs. 9%; fourth graders, 69% vs. 38%). These results indicate that second graders' logarithmic-exponential functions on the 0–1,000 line were not attributable to their not understanding the task.

The excellent fit of the logarithmic model to second graders' performance was inconsistent with two other interpretations. Developmental improvements in ability to estimate proportions might account for the improvement with age in the fit of the linear model, but could not account for the excellent fit of the logarithmic model to the youngest children's estimates. Similarly, the superior fit of the logarithmic function, relative to that of the best-fitting power function, indicates that second graders were not treating the 0–1,000 task as effectively open-ended, a circumstance that has been shown to yield power-function estimation patterns in adults (Banks & Coleman, 1981).

Inclusion of a subset of seven numbers on both the 0–100 and 0–1,000 tasks allowed examination of the third hypothesis: that the same child can represent the same number differently depending on the numerical context. If the numerical context influences the representation that is applied, then the function relating numbers below 100 to positions on number lines may well be logarithmic in the 0–1,000 context, even though it is linear in the 0–100 context.

This turned out to be clearly the case (Fig. 4). For the seven numbers presented in both contexts, the linear function fit median estimates much better in the NP 100 than in the NP 1,000 context (R^2 = .96 vs. .68), $t(6) = 3.06$, $p < .05$. Complementarily, on the NP 1,000 task, the logarithmic function fit the second graders' median estimates for these same seven numbers better than the linear function (R^2 = .89 vs. .68), $t(6) = 2.02$, $p < .10$. Thus, the children had difficulty applying a linear representation in the context of a large scale, rather than an absolute inability to form a linear representation.

The data also allowed us to examine the processes through which adults and sixth graders generated linearly increasing estimates. Participants consistently said they did not know how they generated their responses, but data on the variability of estimates for each number al-

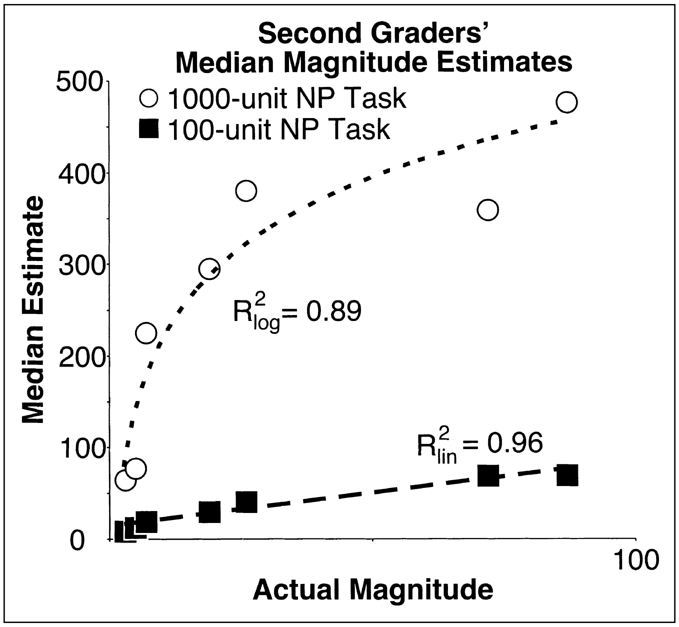

Fig. 4. Second graders' median estimates on the number-to-position (NP) task for the 0–100 and 0–1,000 number lines for numbers presented on both tasks. Also shown are the best-fitting functions—linear for the 0–100 number line and logarithmic for the 0–1,000 number line—and the percentage of variance accounted for by those functions.

lowed tests of several strategies that they might have used implicitly. We initially tested three models—accumulator, pure proportions, and landmark-based proportions. As noted earlier, the accumulator model predicted that the variability of estimates would increase linearly with the size of the number being estimated. The pure proportionality model, in which participants represent each number and spatial position as a proportion of the number line, predicted no systematic pattern of differences among numbers in the variability of estimates. The landmark-based proportionality model predicted that participants would implicitly divide the number line at specific cut points, which they would use as references to guide estimates; the further a number from the closest reference point, the more variable estimates of it would be. On the basis of both intuition and previous findings that division of numbers and quantities into four subgroups is common among both children and adults (Chi & Klahr, 1975; Siegler & McGilly, 1989; Siegler & Robinson, 1982), we hypothesized that participants who generated linear estimation patterns on the 0–1,000 scale did so by dividing the number line into quarters, with divisions at 250, 500, and 750. Thus, estimates for numbers such as 230 and 790, which are near the landmarks 250 and 750, respectively, would be less variable than estimates for numbers such as 390 or 86, which are further from any landmark.

We examined the accuracy of each model's predictions of the variability of estimates for sixth graders and adults, almost all of whom generated linearly increasing estimates on the 0–1,000 number line. The landmark-based proportionality model accounted for significant variance on all four tests (R^2s = .73, .49, .39, and .60 for sixth graders on the NP 1,000 and PN 1,000 tasks and adults on the NP 1,000 and PN 1,000 tasks, respectively). The landmark-based proportionality strategy appeared to be a specific means of implementing the linear representation, rather than being inherent to the experimental task. The same landmark-based proportionality model that fit the estimates of the sixth graders and adults, who consistently generated linearly increasing estimates, did not fit the estimates of second graders, who did not generate linearly increasing estimates. Moreover, the variability of adults' estimates on each number correlated highly with the variability found for sixth graders, who, like adults, generated linear patterns of estimates, *r*s(10) = .75 and .78 on the NP 1,000 and PN 1,000 tasks; in contrast, adults' variability did not correlate with that of second graders, who did not generate linear patterns, *r*s = .30 and .42.

The alternative models fared less well. The accumulator model was clearly incorrect; on the NP 1,000 task, numerical magnitude accounted for less than 5% of variance in the variability of estimates for both adults and sixth graders. The pure proportionality model was equally incorrect. It predicted no systematic pattern of variability differences, whereas the variability of adults' and sixth graders' estimates did vary systematically with distance of the estimated number from a landmark. Moreover, tests of alternative landmark models that divided the number line into thirds, fifths, and tenths showed that each of these models accounted for significant variance in only one of the four analyses of the variability of sixth graders' and adults' estimates (a different analysis in each case). Additional tests are clearly needed, and the number of landmarks may vary with task characteristics, but these data indicate that relying on subjective landmarks is one effective strategy for generating linearly increasing estimates.

In summary, the results of the study support all three of the main hypotheses. They show that individual children know and use multiple representations of numerical quantity; that with development, children rely increasingly on formally appropriate, linear representations rather than intuitive, logarithmic ones; and that the same integers can elicit either a logarithmic or a linear pattern of estimates, depending on the numerical context. Given that adults generate logarithmic and decelerated power functions on other numerical tasks (Banks & Coleman, 1981; Banks & Hill, 1974) and linear functions on the present task, they too appear to utilize multiple numerical representations. The results also show that both the logarithmic-ruler and the accumulator models are too simple to adequately characterize children's and adults' numerical representations. The logarithmic-ruler model accurately characterized second and fourth graders' median estimates, but not those of sixth graders or adults. The accumulator model fit the median estimates of sixth graders and adults but not those of second and fourth graders, and also did not accurately predict the relation between medians and variances of estimates at any age. More generally, it seems unlikely that any model that posits that people use a single, unvarying representation of numerical quantity could fit the totality of the data. Instead, accurate models require the recognition that over a wide age range, people possess multiple numerical representations, with choices among representations changing with age and experience.

Acknowledgments—This research was supported by Grants HD19011 and MH19102 from the National Institutes of Health. We thank Jean Moroney, who made numerous important contributions to the research, and the children and teachers of Mt. Lebanon School District, Pennsylvania.

REFERENCES

Banks, W.P., & Coleman, M.J. (1981). Two subjective scales of numbers. *Perception & Psychophysics, 29*, 95–105.

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

Brannon, E.M., Wusthoff, C.J., Gallistel, C.R., & Gibbon, J. (2001). Numerical subtraction in the pigeon: Evidence for a linear subjective number scale. *Psychological Science, 12*, 238–243.

Chi, M.T., & Klahr, D. (1975). Span and rate of apprehension in children and adults. *Journal of Experimental Child Psychology, 19*, 434–439.

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

Dehaene, S. (2001). Subtracting pigeons: Logarithmic or linear? *Psychological Science, 12*, 244–246.

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

Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is numerical comparison digital: Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 626–641.

Gallistel, C.R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences, 4*, 59–65.

Gibbon, J., & Church, R.M. (1981). Time left: Linear versus logarithmic subjective time. *Journal of the Experimental Analysis of Behavior, 7*, 87–107.

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.

Hiebert, J., & Wearne, D. (1986). Procedures over concepts: The acquisition of decimal number knowledge. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 199–223). Hillsdale, NJ: Erlbaum.

Joram, E., Subrahmanyam, K., & Gelman, R. (1998). Measurement estimation: Learning to map the route from number to quantity and back. *Review of Educational Research, 68*, 413–449.

Keil, F.C. (1989). *Concepts, kinds, and cognitive development.* Cambridge, MA: MIT Press.

Moyer, R.S., & Landauer, T.K. (1967). Time required for judgments of numerical inequality. *Nature, 215*, 1519–1520.

Sekuler, R., & Mierkiewicz, D. (1977). Children's judgments of numerical inequality. *Child Development, 48*, 630–633.

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

Siegler, R.S., & McGilly, K. (1989). Strategy choices in children's time-telling. In I. Levin & D. Zakay (Eds.), *Time and human cognition: A life span perspective* (pp. 185–218). Amsterdam: Elsevier Science.

Siegler, R.S., & Robinson, M. (1982). The development of numerical understandings. In H.W. Reese & L.P. Lipsitt (Eds.), *Advances in child development and behavior: Vol. 16* (pp. 241–312). New York: Academic Press.

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

Starkey, P., & Cooper, R.G. (1980). Perception of numbers by human infants. *Science, 210*, 1033–1035.

Washburn, D.A., & Rumbaugh, D.M. (1991). Ordinal judgments of numerical symbols by macaques (*Macaca mulatta*). *Psychological Science, 2*, 190–193.

Whalen, J., Gallistel, C.R., & Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science, 10*, 130–137.

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

(Received 2/22/02; Revision accepted 8/21/02)