Transitions in Strategy Choices

Robert S. Siegler

Carnegie-Mellon University

Traditionally, transitions in children's thinking have been characterized in elegantly simple ways. For example, 5-year-olds have been said to understand conservation in terms of a single dimension, the height of the liquid columns in the two glasses, whereas 7-year-olds have been said to understand it in terms of transformations. Similarly, first and second graders have been said to add small numbers by counting up from the larger addend, whereas those beyond third grade have been said to solve such problems by retrieving answers from memory (e. g., Groen & Parkman, 1972).

These models of transitions are clean and appealing, but they also are too simple to capture many of the most interesting changes in children's thinking. On a variety of problems, individual children know and use multiple strategies, not just one. They choose among these strategies in ways that produce adaptive combinations of strategy use at any one time and that produce adaptive changes in strategy use over time. They also construct new strategies, which gain a niche among the existing ones, and change the situations in which different strategies are used. Concurrent with all of these changes in strategy use, children become faster and more accurate in executing each strategy and in solving all problems.

This article is organized into four sections. The first describes some of the evidence that children actually use multiple strategies in situations where they previously have been presumed to just use one. The second describes a model of how children choose strategies at any particular point in development. The third describes the part of the model that deals with how children change strategies over time. The fourth describes how children construct new strategies.

Evidence that Children Use Diverse Strategies

Here and in most of the sections that follow, I describe research on 4- to 10-year-olds' strategy choices in arithmetic. My colleagues and I have obtained similar findings in a number of other areas: word identification, time-telling, spelling, and serial recall (Siegler, 1986; in press-a). I focus on the arithmetic research here because the findings are representative of those we have obtained in other domains, and because the models of transition processes have progressed the furthest.

For the past 15 years, the min model has been widely accepted as accurately describing the way that first and second graders solve addition problems. In this model, children consistently solve addition problems by counting up from the larger addend the number of times indicated by the smaller addend. For example, on 3+6, they would think "6, 7, 8, 9". The min model predicts that solution times on each problem will be a linear function of the smaller addend, because the smaller addend indicates the amount of counting-on from the larger number that needs to be done to solve the problem.

This prediction has proved accurate for both groups of children and individuals, in both Europe and North America, and in both standard and special education settings.

Despite all this support, the min model is wrong. Siegler (1987) examined young children's simple addition, using both the usual solution-time measures and children's verbal reports. The results were striking. When data were averaged over all trials (and over all strategies), as in earlier studies, the results replicated the previous finding that solution times were a linear function of the smaller addend. If these analyses had been the only ones conducted, the usual conclusion would have been reached, namely that first and second graders consistently use the min strategy to add.

However, the children's verbal reports suggested a quite different picture. The min strategy was but one of five approaches that they reported using. This reporting of diverse strategies characterized individual as well as group performance; most children reported using at least three approaches. Children reported using the min strategy on only 36% of trials.

Dividing the solution time data according to what strategy children said they had used on that trial lent considerable credence to the children's verbal reports. On trials where they reported using the min strategy, the min model was an even better predictor of solution times than in past studies or in the Siegler (1987) data set as a whole; it accounted for 86% of the variance in solution times. In contrast, on trials where they reported using one of the other strategies, the min model was never a good predictor of performance, either in absolute terms or relative to other predictors. It never accounted for as much as 40% of the variance. A variety of measures converged on the conclusion that children used the five strategies that they reported using, and that they employed them on those trials where they said they had.

Models of transitions can be no better than the characterizations of early and later knowledge states that they are attempting to connect. In arithmetic and many other domains, adequately characterizing these knowledge states demands recognition of children's use of diverse strategies over extended periods of development.

A Model of How Children Choose Which Strategy to Use

Once we recognize that children use diverse strategies to solve many problems, it becomes essential to identify how they choose among them. For the past few years, my colleagues and I have been developing a model of how children choose among their diverse strategies. The model has been implemented in detail (as a running computer simulation) for addition, subtraction, and multiplication. In all of these areas, the simulations produce strategy choices at any given time, changes in strategy use over time, and improvements in accuracy and speed much like those of the children we have observed.

The current version of the simulation, which I will describe here, is a more general version of the addition simulation described by Siegler and

Shrager (1984). Like the previous version, it includes a representation of knowledge and a process that operates on that representation to produce performance and learning.

First consider the representation. Children are hypothesized to have knowledge of problems, of strategies, and of the interaction between problems and strategies. Their knowledge of problems involves associations between each problem and possible answers to that problem, both correct and incorrect. For example, 5+3 would be associated not only with 8 but also with 6, 7, and 9. These representations of knowledge of each problem can be classified along a dimension of the peakedness of their distribution of associations. In a peaked distribution, most associative strength is concentrated in the correct answer. At the other extreme, in a flat distribution, associative strength is dispersed among several answers, with none of them forming a strong peak. For example, in Figure 1, the associative strengths for answers to 2+1 form a peaked distribution (with the associative strength for 3 at the peak) and those for 3+5 form a flat distribution.

The representation also includes knowledge about strategies. Each time a strategy is used, the simulation gains information about its speed and accuracy. This information generates a strength for each strategy, both in general and on particular problems. The strategies modeled in the current version of the addition simulation are the three most common approaches that young children use: counting from one, the min strategy, and retrieval.

One further feature of the representation should be mentioned. Newly generated strategies possess "novelty points" that temporarily add to their strength and thus allow them to be tried even when they have little or no track record. The strength conferred by these novelty points is gradually lost as experience with the strategy provides an increasingly informative data base about it. This feature was motivated by the view that people are often interested in exercising newly developed cognitive capabilities (Piaget, 1952), and by the realization that without a track record, a newly-developed strategy would be unlikely to be chosen.

Now consider the process that operates on this representation to produce performance. First, the process chooses a strategy. The probability of a

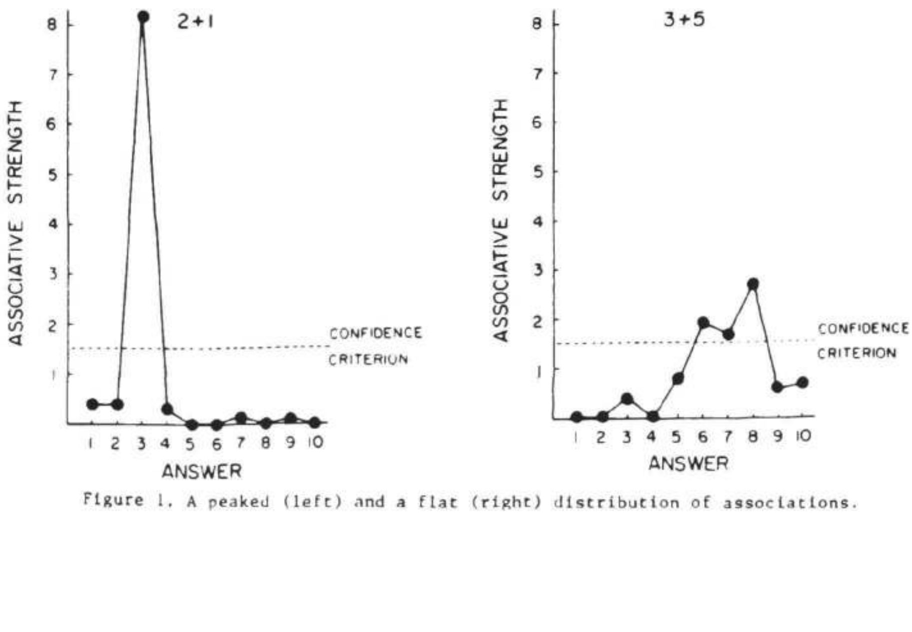

Figure 1. A peaked (left) and a flat (right) distribution of associations.

given strategy being chosen is proportional to its strength relative to the strength of all strategies. Strength of a strategy on a problem is a joint function of the local value of the strategy (how well it has done on that problem in the past) and of its global value (how well it has done across all problems). On problems never previously encountered, the global value of the strategy is the sole determinant of its strength. Thus, the stronger a strategy in general and on the particular problem that is posed, the more likely that it will be chosen for use on that problem.

If a strategy other than retrieval is chosen, that strategy is executed. If retrieval is chosen, the simulation retrieves a specific answer (e. g., 4) from the problem's distribution of associations (Figure 1). The probability of any given answer being retrieved is proportional to that answer's associative strength relative to the strength of all answers to the problem. Thus, in Figure 1, the connection between "2+1" and "3" has a strength of .80, the strength of connections between "2+1" and all answers is 1.00, so the probability of retrieving "3" is 80%. If the associative strength of whichever answer is chosen exceeds the confidence criterion (a threshold for stating a retrieved answer), the simulation states that answer. Otherwise, the simulation again chooses a strategy with probability proportional to the strength of that strategy relative to those of all strategies. The process continues until a strategy is chosen and an answer stated.

The simulation generates patterns of accuracy, solution times, and strategy use much like those of children. For example, it uses the min strategy most often on problems where the smaller of the two addends is very small and where the difference between the two addends is quite large. Siegler (1987) found the same pattern in kindergarteners', first graders', and second graders' performance. Also as with children, the simulation uses retrieval most often on problems where both addends are small and uses counting-from-one primarily on problems where both addends are large. Relative problem difficulty and particular errors that the simulation makes also parallel those of children. The reason lies in the simulation's learning mechanism, which is described in the next section.

Transitions in Strategy Use Over Time

The simulation learns a great deal through its experience with strategies and problems. As it gains experience, it produces faster and more accurate performance, more frequent use of retrieval, less frequent use of counting from one, and closer fitting of when strategies are used to their advantages and disadvantages on each problem. Such learning is not produced by any explicit, metacognitive governmental process, but rather through the operation of the above-described program together with a simple learning mechanism: children associate answers that they state with the problem on which they state them, and associate each strategy with the speed and accuracy that the strategy has produced on each problem and over all problems.

The way that this learning mechanism operates can be illustrated in the context of why some strategies are assigned to some problems more than others. Consider two problems, 9+1 and 5+5. Kindergarteners and first graders use the min strategy considerably more often on 9+1, yet use counting-from-one more often on 5+5. The simulation generates similar behavior, and illustrates how such a pattern might emerge. On 9+1, the min strategy has a very large advantage in both speed and accuracy over the count-all strategy. It requires only 1/10 as many counts. In contrast, the numbers of counts required to execute the two strategies are more comparable on 5+5, where the min strategy requires 1/2 as many counts. If the number of counts were the only consideration, children might be expected to consistently use the min strategy on both problems (and on all problems) from the time they learned it. However, for any given number of counts, counting-on from an arbitrary number is considerably more difficult for young children (in terms of time and errors per count) than counting from one Fuson and Richards, 1982). The simulation's probability of erring on each count, and its time per count, reflect this greater difficulty of counting-on from a number larger than one. Thus, the simulation learns that although the min strategy is generally more effective, there are some problems, such as 5+5, where counting-from-one works better. This leads to counting-from-one being the most frequent strategy on such problems for awhile. It eventually is overtaken by retrieval, however, as the associative strength of the correct answer becomes sufficiently great that it is likely to be retrieved and stated.

The influence of performance on learning also is reflected in relative problem difficulty and in the particular errors that children make. Early in learning, children most often use backup strategies (such as counting-from-one and the min strategy). Early patterns of difficulty in executing such backup strategies seem to influence later patterns of retrieval difficulty and particular errors that are made. For example, in multiplication, the most common backup strategy is repeated addition. Repeated addition generates two main types of errors: answers in which one multiplicand is added too many or too few times (e. g., on 8X4, adding 7 or 9 4's, and getting 28 or 36) and small addition errors (e. g., adding 8 4's and getting 33). These are the same types of errors that children make most often when retrieving answers and that adults make under time pressure. Similarly, third and fourth graders' probability of correctly executing repeated addition on simple multiplication problems is highly correlated with their probability of retrieving correctly (Siegler, in press-a). The same relation between difficulty of solving problems via backup strategies and via retrieval has been found in addition and subtraction (Siegler, 1986).

The simulation's learning mechanism also produces parallels between backup strategy and retrieval performance. Problems that are more difficult to solve early in learning via backup strategies become more difficult to solve later via retrieval. When the simulation errs in using backup strategies, the result is less opportunity to associate the problem with the correct answer and more opportunity to associate incorrect answers with it.

Backup strategies also influence their own future use, in a somewhat ironic way. The more accurately that they are executed within the simulation, the sooner they stop being used. The reason is that accurate execution of backup strategies leads to peaked distributions of associations between answers and each problem, which in turn leads to retrieval, inherently the fastest strategy, also generating accurate performance and therefore being used increasingly often. Thus, backup strategies contribute to the transition process in ways that lead to their own demise. Consistent with this view, the accuracy with which children execute backup strategies is quite strongly correlated with how often they are able to correctly retrieve answers in addition, subtraction, and word identification (Siegler, in press-b).

Acquisition of New Strategies

The computer simulations generate a fair range of transitions in children's arithmetic, but by no means all of them. Perhaps the most conspicuous gap is in the account of how new strategies are acquired. To learn more about this issue, Eric Jenkins, a graduate student at CMU, and I are currently conducting a longitudinal study of acquisition of the min strategy. Past studies indicate that children ordinarily acquire this strategy at age 5 or 6, and that 4-year-olds can learn it if given extensive addition experience.

To study the acquisition process, Jenkins and I pretested a group of 4- and 5-year-olds to identify children who gave no evidence of prior knowledge of the min strategy. Once the children were identified, we presented them 7 problems daily for about 30 sessions. After each problem, we asked the child how he or she had solved the problem. The verbal statement, together with the videocassette of the child's overt behavior while solving the problem, became our guide for identifying the child's strategy on each trial. This gave us a way of identifying the exact trial on which the child discovered the min strategy, and thus to analyze what led up to the discovery and how the strategy, once discovered, was extended to new problems.

The min strategy was discovered by 7 of the 9 children in the experiment. Some children constructed the strategy within the first 5 sessions; others did so between the 25th and 30th sessions. For some children, invention of the new strategy was accompanied by conscious appreciation that they were doing something new and that the new way of adding was more efficient. Other children were unaware that they were doing anything different, even saying that they had counted from one when they had audibly began counting at the larger addend.

The most striking finding of the study involved a condition that seemed to strongly promote both discovery of the min strategy among children who had not already discovered it and increased use of the strategy among those who had. This condition involved presentation of problems that would be very difficult to solve by means of the prior counting-from-one strategy, but that would be quite easy to solve via the min strategy. These problems involved adding very large and very small numbers, scuh as 24+3 and 2+23. Within two sessions of the introduction of these problems, use of the min strategy increased from 15% to 65% of trials on which any type of counting was used. The progress was maintained after the demanding problems were no longer given and more standard problems (e. g., 7+4) were substituted. Thus, children met the challenge of the difficult problems by constructing a new strategy, and then continued to use the new strategy on other problems.

Looking at the present research as a whole, perhaps the most striking characteristic of cognitive-developmental transitions is their self-regulating quality. The transitions are self-regulating in at least three senses. First, children's choices of strategies adapt to changing circumstances. As they gain experience with a strategy, they use it increasingly often on those problems where its advantages relative to other strategies are greatest. Second, children's strategy choices have built into them a kind of self-righting capability, an ability to recover from errors and initial unfavorable experiences. The heavy use of backup strategies early in learning confers this type of stability. Illustratively, the simulation of multiplication reported in Siegler (in press-a) erred on its first 4 answers on 8X9 and on 8 of its first 10. Yet by the end of the learning phase, the simulation was advancing the correct answer on 99% of trials. The reason was that over trials, the backup strategies produced 72 more often than any other answer, which led to its associative strength increasing and therefore to its being retrieved and stated increasingly often.

Third, when existing procedures prove inadequate, children are especially likely to create new strategies that can overcome the difficulties. The present models produce the first two types of self-regulation; I hope soon to incorporate mechanisms into the models, perhaps akin to those in Newell's Soar or VanLehn's Sierra models, that produce the third type of self-regulation as well.

References

Fuson, K. C. & Richards, J. (1982). The acquisition and elaboration of the number sequence. In C. Brainerd (Ed.) Progress in Cognitive Developmental Research: Vol. 1. New York: Springer-Verlag, 1982.

Groen, G. J., & Parkman, J. M. (1972) A chronometric analysis of simple addition. Psychological Review, 79, 329-343.

Piaget, J. (1952). The origins of intelligence in children. New York: International Universities Press.

Siegler, R. S. (1986). Unities across domains in children's strategy choices. In M. Perlmutter (Ed.) Minnesota Symposium on Child Psychology, Vol. 19, (pp. 1-48), Hillsdale, NJ: Erlbaum.

Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. Journal of Experimental Psychology: General, 116, 250-264.

Siegler, R. S. (In press-a). Strategy choice procedures and the development of multiplication skill. Journal of Experimental Psychology: General.

Siegler, R. S. (In press-b). Individual differences in strategy choices: Good students, not-so-good students, and perfectionists. Child Development.

Siegler, R. S., & Shrager, J. (1984). A model of strategy choice. In C. Sophian (Ed.), Origins of cognitive skills (pp. 229-293). Hillsdale, NJ: Erlbaum.