Chapter 2
Rightstart: Providing the Central Conceptual Prerequisites for First Formal Learning of Arithmetic to Students at Risk for School Failure

Sharon A. Griffin, Robbie Case, and Robert S. Siegler

A desire to improve mathematics education is not a new phenomenon in the history of U.S. schooling. Although it has been expressed for decades, in various quarters and with varying levels of intensity (Case and Garret 1992), it assumed significant and widespread proportions in the mid-1980s when the results of international comparisons were widely circulated.

These findings suggested that American children were not acquiring the mathematical knowledge possessed by their Asian peers or the skills needed to succeed in a technologically advanced society (NAEP 1983, 1988). This knowledge gap was found to be minimal in grade one, to become increasingly pronounced as schooling progressed, and was apparent on a wide array of measures that tapped conceptual understanding, as well as knowledge of facts and procedures (Steven- son, Lee, and Stigler 1986; Stigler and Perry 1988). Of significance to the work described in this chapter, this knowledge gap was most pronounced in the performance of American children living in economically deprived inner-city communities (Case 1975; Saxe, Guber- man, and Gearhart 1989).

A dominant explanation that was offered for these findings was that the formal learning opportunities provided in American schools are divorced from children’s intuitive, informal understandings (see, for example, Ginsburg and Ruset 1981; Hiebert 1986). Armed with re- search findings that accumulated in the 1980s, and that documented the manner in which children’s intuitive knowledge of number develops across the school-age years (e.g., Resnick 1989; Siegler and Robin- son 1982; Case 1985), the response from the educational community was swift and sure. The National Council of Teachers of Mathematics published a mandate (NCTM 1989b) requiring that mathematics in- struction be grounded in children’s intuitive knowledge. Several pro- grams that implement this mandate are currently being developed (see, for example, University of Chicago School of Mathematics Project
1990; Resnick, et al. in press; Cobb et al. 1991; Feinnema and Carpenter 1989; Kamil 1985). These solutions promise to do much to reduce the knowledge gap described above.

However, in recent cognitive developmental work we have conducted, another, potentially more fundamental, source of the knowledge gap became apparent. Although many children start school with a well-developed intuitive understanding of number (e.g., Hiebert 1986; Case 1985; Siegler and Robinson 1982), not all children do so.

In particular, when tests of conceptual knowledge were administered to groups of kindergarten children attending schools in low-income inner-city communities, a significant number were unable to demonstrate the knowledge possessed by their middle-income peers (Case and Griffin 1990; Griffin, Case, and Capodilupo in press). When tests of procedural knowledge were administered to groups of kindergarten and first grade children in similar communities, a significant number used strategies that were "non-adaptive," that appeared to prevent the development of more adaptive strategies, and that were never found in the strategy choices of their middle-income peers (Siegler in press). If children start school without the intuitive knowledge that is explicitly assumed by the mathematics programs being developed for the 1990s, the risk for school failure may continue, in spite of the excellence of the new programs.

The Rightstart program was designed to remedy this knowledge gap at its source, at the very start of formal schooling. By developing a program to teach the central conceptual prerequisites for first formal learning of arithmetic (i.e., the intuitive knowledge on which success in addition and subtraction depends), we hoped to ensure that all children who had not already acquired the critical knowledge had an opportunity to do so before they started first grade. In the following sections, we describe the research evidence that suggested a need for the Rightstart program, the theoretical framework that was used to develop it, the program itself, and the results that were obtained when this program was implemented over a three-year period.

Research on Young Children’s Strategy Use

Information processing research has revealed a great deal about how young children solve arithmetic problems. They appear to use a variety of strategies to do so, including counting from one, counting from the larger addend, decomposing a single problem into two or more simpler ones, and retrieving answers from memory (Ruson 1982; Green and Parkman 1972; Siegler in press). Choosing wisely among these strategies offers children the opportunity to fit the characteristics
of the strategy to the demands of the task. However, not all children make wise choices. Some use fast and easy approaches, even when these approaches have little likelihood of success; others use unnecessarily slow and time-consuming approaches, even when they could succeed with faster and easier ones.

These individual differences in children's strategy choices are apparent as early as first grade. In a series of studies conducted with children from middle-income communities, Siegler (1986) examined individual differences in 6-year-olds' strategy choices in addition and subtraction, using the model he had developed and validated earlier with Shrager (Siegler and Shrager 1984). What he found was that children could be clustered into three distinct groups. "Good students" had a core set of problems they knew "by heart," and for which they would retrieve the answers from memory. For problems they did not know as well they also had a set of efficient "backup strategies," which would allow them to compute the answers from first principles (e.g., via counting). By contrast, "not-so-good" students were poor both at retrieving answers and at using counting strategies. They also used retrieval on many problems where they had little chance of doing so correctly. Finally, there was a group of "perfectionists," whose knowledge of the basic problems was about the same as that of the good students, and whose backup strategies were just as efficient, but who employed a far stricter criterion for retrieving answers from memory, thus laboriously computing a great many answers for which they knew the answers quite well already.

This analysis of individual differences received a number of types of external validation. For example, good students' and perfectionists' achievement test scores averaged at the 81st and 80th percentiles respectively, whereas the not-so-good students' scores averaged at the 43rd percentile. Further, all of the children who were assigned to learning disabilities classes or held back at the end of the year were in the not-so-good group. Thus, the differences between not-so-good students and the other two groups were evident on standardized tests as well as in the children's strategy choices. Only the strategy-choice test, however, revealed the stylistic differences between the perfectionists and good students. In addition, only the strategy-choice test could specify the precise locus of the poor students' difficulties.

Recently, Siegler and Kerkman (Siegler in press; Kerkman and Siegler in press) examined whether first graders from low-income communities would show similar patterns of strategy use and strategy choice, and similar individual differences in these properties, to those observed with middle-income children. The study exactly paralleled the Siegler (1988) study of individual differences, except that the chil-
dren were examined somewhat later in first grade (April rather than November and December). The reason for the later testing was to achieve greater similarity in absolute levels of performance (it was the individual differences in patterns of performance that were of interest).

As hoped, absolute levels of accuracy, speed, and strategy use of the lower-income first graders were comparable to those of the middle-income children tested five months earlier in the school year. The lower-income children also showed highly adaptive choices among alternative strategies, and the general pattern of strategy use closely resembled that found in the middle-income samples. On more than 90 percent of trials, children used strategies that had been observed relatively often in previous studies. In order of frequency of use, these strategies were retrieval (51 percent of trials), counting from the larger addend (16 percent), counting from one (14 percent), finger recognition (7 percent), guessing (3 percent), and saying, "I don't know" (3 percent).

On the remaining 5 percent of trials, however, the children used two other strategies. One was an approach that had rarely been seen in previous detailed strategy assessment studies (all of which had been conducted with middle-class children); the other strategy of interest was one that had literally never been seen. The more common of the two approaches was counting-on from the smaller addend. On a problem such as 5 + 9 or 9 + 5, the child would count-on from the 5. The less common of the approaches was counting on from neither addend. On 5 + 9 or 9 + 5, the child might count on from 7 or 8.

Analyses of individual children's counting from the smaller addend were especially revealing. One group, including 36 percent of children, counted from the smaller addend on fewer than 10 percent of the trials on which they counted from one of the two addends. Another group, including 11 percent of children, counted from the smaller addend on at least 50 percent of trials on which they counted from one or the other addend. However, the largest group—53 percent—counted from the smaller addend on between 10 percent and 50 percent of such trials. These data demonstrate that counting from the smaller addend could not be attributed to a small number of children engaging consistently in the activity. Rather, it seemed to be something that many of the lower-income children did sometimes. The pattern contrasted with the findings of Siegler (1987) and Siegler and Jenkins (1989), studies that applied the same strategy assessment methodology to middle-income samples. In both of the studies with middle-income samples, counting from the smaller addend was an extremely rare activity. Only one of the 4- and 5-year-olds in the Siegler and Jenkins study and none of the middle-income 6- and 7-year-olds in Siegler
(1987) were observed to use it (versus three of four children from the low-income families in Siegler and Kerkman, in preparation).

The other strategy of interest was counting from neither addend, an approach that had never been observed in any previous studies of children from middle-income backgrounds. Even in Siegler and Jenkins (1989), where 4- and 5-year-olds' very first uses of counting from a number larger than one were focused on, the children never counted from a number different than either addend.

Although counting from the smaller addend fairly often generates correct answers, and counting from neither addend hardly ever does, it seems quite likely that the two strategies reflect different degrees of a similar problem: an early gap between these children's use of a strategy and their understanding of why the strategy works and what goals it achieves. In short, the pattern apparent in these findings suggests that many children in the lower-income groups may not have acquired the intuitive understanding of number that was demonstrated in the strategy use of the middle-income group.

Research on Young Children's Intuitive Knowledge of Number

After reviewing the literature on young children's arithmetic, Resnick (1983) concluded that children represented the addition process in terms of something like a mental number line, and that this number line provided the conceptual underpinning for children's use of arithmetic strategies. According to her analysis, children see the use of a strategy such as counting on their fingers as being justified, because it is like traversing a mental number line and arriving at the answer. The lack of such a representation may be part of the reason that some students have difficulty in learning to add or subtract or are poor at executing these strategies; lack of understanding of why the strategies work may interfere with skilful execution of the strategies. This, in turn, may lead to such strategies not being used very often, and may prevent a core set of problems from being learned "by heart."

In subsequent work, Case, Griffin, and their colleagues (Case and Sandieson 1987; Case and Griffin 1990; Griffin, Case, and Sandieson 1992; Okamoto 1992) developed, tested, and refined a measure to assess children's conceptual knowledge of number. The test was used initially with middle-income American children, and normed for the age levels of 4, 6, 8, and 10 years. The developmental theory that guided construction of these measures is described in the following section but one set of findings is relevant to the present discussion.

When the number knowledge test (see Table 2.1) was first administered to a group of low-income 5- to 6-year-olds, they were found to
Table 2.1
Number knowledge test

Preliminary. Let’s see if you can count from 1 to 10. Go ahead.

Level 0: (4 years)
1. (Present mixed array of 3 red and 4 blue poker ships.) Count the blue chips and tell me how many there are.
2. I’m going to give you 1 candy and then I’m going to give you 2 more. (Do so.) How many do you have altogether?
3. (Show stacks of 5 and 2 poker ships, same color.) Which pile has more?
4. (Present mixed array of 7 circles and 8 triangles, same color.) Count the triangles and tell me how many there are.

Level 1: (6 years)
1. If you had 4 chocolates and someone gave you 3 more, how many chocolates would you have altogether?
2. What number comes right after 7?
3. What number comes two numbers after 7?
   4a. Which is bigger: 5 or 4?
   4b. Which is bigger: 7 or 9?
   5a. Which is smaller: 5 or 6?
   5b. Which is smaller: 5 or 7?

   6a. (Present visual array.) Which number is closer to 5: 6 or 2?
   6b. (Present visual array.) Which number is closer to 7, 4 or 9?

Level 2: (8 years)
1a. Which is bigger: 69 or 71?
1b. Which is bigger: 32 or 28?
2a. Which is smaller: 27 or 32?
2b. Which is smaller: 51 or 39?

3a. (Show visual array.) Which number is closer to 21, 25 or 18?
3b. (Show visual array.) Which number is closer to 28, 31 or 24?

perform much like middle-income 3- to 4-year-olds. Unlike middle-income 5- and 6-year-olds, a significant number of these children were unable to tell which of two numbers is bigger (e.g., 6 or 8), or which number is closer to 5, 6 or 2. They were also unable to answer the verbally presented problem involving the receipt of four objects followed by three more objects.

While the difficulties on the first two items suggest an absence of the general conceptual knowledge inherent in a mental number line representation, the latter difficulty indicates the absence of that aspect of this knowledge on which the solving of first grade addition and subtraction problems is most directly dependent. Most middle-income kindergarten children already knew how to solve simple addition problems when these were presented orally. Thus, the only training
they really needed was in conventional numerical notation, such as 4 + 3 = ?. By contrast, the low-income group often gave answers that seemed like wild guesses (e.g., 4 + 3 = 13) or simple associations based on their knowledge of counting (e.g., 4 + 3 = 5). In short, they appeared to be missing precisely those conceptual prerequisites on which success in arithmetic depends.

On the basis of the findings described thus far and the theoretical formulations described in the following section, the Rightstart project was launched. Subsequent findings as the project progressed, and as the number knowledge test was administered to an increasing number of children from low-income communities, provided evidence that the knowledge gap identified in the first study was not an isolated phenomenon. It was apparent in each of the studies reported in this chapter, in the performance of kindergarten children attending schools in Canada, California, and Massachusetts. An overview of these findings for selected items on the Number Knowledge test is provided in table 2.2.

Theoretical Framework

In the past decade or so, a number of neo-Piagetian theorists have proposed that a major reorganization occurs in children’s thought at 5-6 years, about the age when arithmetic is normally introduced in schools (Case 1985; Fischer 1980; Halford 1982).
Our own efforts to investigate this reorganization began with Siegler's (1976) observation that 6-year-olds use a quantitative rule for solving balance beam problems ("Pick the side that has the greater number of weights as the one that will go down"). Four-year-olds cannot usually solve balance beam problems that demand a role of this sort. However, they can solve problems that are solvable with a qualitative version of this rule, that is, problems where one side has a very large weight and the other side a very small one (Liu 1981). Four-year-olds also can solve problems where their only task is to count small arrays of objects. Given this pattern of strengths and weaknesses, Case and Sandieson (1987) proposed that what 4-year-olds lack is the ability to coordinate these two schemes into what they termed a "dimensional structure", that is, a structure in which properties such as weight are represented as quantitative dimensions with two poles and a continuum of values in between.

In a series of studies conducted in the 1980s, this difference between 4- and 6-year-olds was found to be quite general. It was apparent on tasks of moral reasoning (Damen's Distributive Justice task), tasks of social reasoning (Marini's Birthday Party task), and tasks of scientific reasoning (Piaget's Projection of Shadows task). Children's performance across all of these tasks was also found to be highly consistent at the age-levels of 4, 6, 8, and 10 years (Marini 1992). To account for these findings, we proposed that each task had some (often implicit) quantitative component and that children's performance on all these tasks, drawn from distinct content domains could be explained by their growing understanding of number (Case and Griffin 1990; Griffin, Case, and Sandieson 1992). Using postulates of neo-Piagetian theory (Case 1985, 1992), we modeled the quantitative understandings that appeared to be central for successful performance at four age-levels as depicted in figure 2.1.

What this figure is meant to suggest is that, at the age of 4 years, children tend to represent all possible variables in a global or polar fashion, so that they can make mappings of the sort: "Big things are worth a lot; little things are worth a little." At the age of 6 years, children tend to represent variables in a continuous fashion (i.e., as having two poles and a number of points in between). Moreover, they realize that these points can be treated as lying along a mental number line, such that values which have a higher numeric value also have a higher real value associated with them. At the age of 8 years, children can think in terms of two independent quantitative variables (e.g., hours and minutes on a clock; dollars and cents), but cannot yet make successful comparisons between variations along each. Finally, at the age of 10 years, children can make these sorts of
comparisons, by thinking in terms of the "trade-offs" between two quantitative variables.

To test this model, two additional instruments were constructed, a time-telling test and a money-handling test. Items included in these tests were drawn from the domain of everyday experience and were assigned to four age-levels in accordance with the model's predictions. When these tests were administered to populations of children from middle-income communities, the predictions were found to be highly accurate. At four age-levels (4, 6, 8, and 10 years), children's performance was consistent with the model and consistent with the pattern found in previous studies (Griffin, Case, and Sandieson 1992). Given the generality of the findings, we proposed that the numerical understandings implied in the model were powerful organizing schemata, "central conceptual structures," that mediate performance on a wide variety of tasks with some quantitative component (Case and Griffin 1990; Case 1992).

The number knowledge test that was described in the previous section was developed to assess the knowledge implied in these structures. When the findings suggested that kindergarten children in one at-risk community could not pass the 6-year-old level of the test, we
Central conceptual structure hypothesized to underlie six-year-old's early understanding of mathematics (dotted lines indicate "optional"—i.e., non-universal—notational knowledge).
began to speculate that the reason some American children do poorly in school is that they haven’t acquired the central conceptual structure on which success in arithmetic depends. If we could develop a program to teach this structure, we might give children who did not already possess it, a solid conceptual foundation for first formal learning of arithmetic as well as a foundation for higher-order mathematical understandings. The Rightstart program was designed to serve this purpose.

To teach this structure, we required a more detailed model of the knowledge implied in a single mental number line. Relying on a rational analysis to identify central components of this structure, results of the number knowledge test, findings in the mathematics research literature (e.g., Resnick 1983; Gelman and Gallistel 1978) which helped us identify knowledge components that serve as precursors to a full understanding, and in-depth interviews with children, we proposed the elaborated structure depicted in figure 2.2.

The understandings implied in this figure and presumed to be conceptual prerequisites for all arithmetic operations with single-digit numbers, can be described as follows: (a) a knowledge of the number sequence from 1 to 10, and an awareness of each number’s position in the sequence; (b) a knowledge of the one-to-one manner in which this sequence is mapped on to objects when counting; (c) an understanding of the cardinal value of each number (i.e., that when touching the third object and saying “3,” one has formed a set whose size is indicated by this number); (d) an understanding of the generative rule which relates adjacent cardinal values (e.g., that 3 represents a set that’s just like 2 except that 1 object has been added, or that 3 represents a set that’s just like 4 except that 1 has been subtracted); and (e) an understanding of the consequence of this fact: namely, that each successive number represents a set which contains more objects, and thus has a greater value along any particular dimension. The first three elements specified above are known to develop in the preschool years (Gelman and Gallistel 1978).

To summarize, the theoretical assumptions on which the Rightstart program was based are as follows: (1) in the course of their preschool experience, many (but not all) children develop powerful organizing schemata that are central to their understanding of the school tasks that they encounter in subsequent years. These schemata may be called “central conceptual structures” (Case and Griffin 1990). (2) The central conceptual structure on which early addition and subtraction are dependent is one for conceptualizing the world in terms of quantitative dimensions. Once this structure is in place, children see the world as comprised of dimensions having two poles, and a large number of
points between these poles, whose relative magnitude can be indexed by the number system (see figure 2.2). (3) In comparison to their peers, children who are at risk for early school failure in math are less likely to have developed this structure by the time they enter first grade. (4) If these children are provided with experience that enables them to develop this structure—before they enter grade one—their first learning of addition and subtraction should improve considerably.

The Rightstart Program

The curriculum we developed was more limited in scope than many kindergarten programs. It was designed to teach the specific set of nodes and relations specified in the structure depicted in figure 2.2 (i.e., the cognitive structure underlying a 6-year-old's numerical understanding) and to give children multiple opportunities to assemble these components into a well-consolidated whole.

To teach this structure, we developed a series of thirty interactive games (some with several variations) that provided hands-on opportunities for children to construct, and to consolidate, the understandings depicted in each set of nodes and relations within the structure. Sufficient games were developed so that each knowledge component included in the structure could be explicitly targeted in one set of lessons. In addition, each game implicitly targeted several other components of the structure as well, to provide opportunities for multiple levels of learning and to ensure that the integrated structure we were attempting to teach was adequately represented in each learning activity.

The games themselves were designed to be played by a group of 4-5 children, with teacher supervision. Activities that could be used with a whole class of children to reinforce the knowledge acquired in particular games were also included in the program, as were Learning Center activities that children could engage in without teacher supervision. To support these games and activities, we selected props that were congruent with the structure we were attempting to teach, and that represented this structure in a variety of ways.

The Number Line Game

The number line game exemplifies our approach. This game is played by a small group of children on a game board that portrays a series of number lines (one for each player) coded in different colors. To play, children roll a die and compute the quantity shown to determine who goes first. The first player then rolls the die, computes the quantity, asks the banker for that many counting chips, places the chips in
sequence along his or her number line while counting, and then moves his or her playing piece along the counting chips (while counting once again) until the playing piece rests on the last chip (which itself rests just below the numeral that corresponds to the quantity rolled and moved). Play then moves to the next player; the first child to reach the winner’s circle (beyond the 70 square) wins the game. Children are asked to watch and listen carefully while others are counting and moving to ensure that no one makes a mistake. When a mistake is made or a computation is challenged, it is typically resolved by recounting and/or group discussion.

These aspects of game play give children repeated exposure to the vertical nodes and relations illustrated in figure 2.3. They were designed to help children consolidate knowledge of the number sequence up and down, the corresponding numerals, the one-to-one manner in which numbers map onto objects when counting in either direction, and the cardinal value of each number.

When children are comfortable with this level of play (i.e., when they can reliably count quantity sets, and match sets to numbers), several variations are introduced. Children are asked to make relative quantity assessments with questions such as "Who is closest to the goal?" and to map these assessments onto the number sequence with questions such as "How do you know?". They are asked to draw chance cards during game play which require that their quantities, and their position on the number line, be incremented or decremented by one. They are asked to use numerosity to make predictions about who is likely to win, or lose, the game. These variations provide plenty of exposure to the horizontal nodes and relations illustrated in the figure. They were designed to help children acquire the increment and decrement rules, and understand that numbers can be used to make relative quantity assessments (e.g., if Maria has 5 and Stephen has 4, then Maria has more/is further ahead/is closer to the goal).

The twenty-nine other games included in the program are distinct from this one, but they also provide opportunities for children to construct and to consolidate the same knowledge structure, at their own pace and in contexts that are highly motivating for a wide variety of children. In contrast to the competitive format of the game just described, over 80 percent of the games included in the module use a cooperative format in which children must work together in pairs, or in a small group, to achieve the goal. Opportunities or requirements to justify and explain a particular quantitative assessment are built into many of the games. Prompts for verbal communication are scripted in the teacher’s manual, in the form of "How do you know?" questions, for several other games.
Finally, the entire set of games is sequenced within the module to provide a natural bridge between children’s entering understanding of number and quantity, and the complete set of understandings implied in the central conceptual structure. In other words, the knowledge components that are targeted early in the instructional sequence are those that children are known to acquire earliest in development. With room for adaptation to accommodate the particular group of learners, the entire sequence was designed to recapitulate the natural developmental progression.

The instructional principles that were used to develop the program are apparent in the above discussion. They can be summarized as follows:

1. Conceptual bridging: The activities included in the module should serve as a “conceptual bridge” between children’s current understanding of number and quantity and the sorts of understanding implied in the “mental number line” structure.
2. Representational congruence and diversity: The props used to accompany these activities should be congruent with the “mental number line” structure and should represent this structure in diverse ways.
3. Multiple levels of understanding: The activities should allow for multiple levels of understanding so children with different entering knowledge, and different learning rates, can all learn something from each activity.
4. Affective engagement: The activities should be affectively, as well as cognitively, engaging.
5. Physical, social, and verbal interaction: The activities should provide opportunities for children to interact with the materials and to use the knowledge they are constructing in a variety of ways (e.g., to solve game problems; to communicate with peers).
6. Developmental sequencing: The activities should be sequenced in their normal order of acquisition.

Program Evaluation

The Rightstart program was used over a period of three years with small groups, or whole classes, of kindergarten children attending schools in Canada, California, and Massachusetts. In four of the studies reported in this chapter, the children were taught in small groups of 4-5 children, for about 20 minutes a day. Most of the teachers were research assistants who were trained to teach the curriculum. In a fifth study (i.e., study 4), the program was taught by a teacher-researcher
to two whole classes of children, with 20 to 25 children in each class. The entire program extended over a 3- to 4-month period in each study.

The children who received the program were all attending kindergarten in inner-city schools with large minority populations. Most of the children came from low-income communities and were drawn from school populations considered to be at risk for school failure. Prior to training, in the middle of their kindergarten year, the vast majority of the children who received the program (in most studies, the majority of children in each kindergarten class) failed the 6-year-old level of the number knowledge test described earlier in this paper. Many of these children also failed items at the 4-year-old level of this test.

In three studies (i.e., studies 1, 2, and 4), we created matched control groups of children, on the basis of number knowledge test scores, age, school placement, and ethnic background. Two control groups were established for study 1, which was conducted in a public school in metropolitan Toronto that drew children, almost exclusively, from a Portuguese immigrant community. Most children included in the sample had received a year of preschool training (i.e., junior kindergarten) and all were proficient in English when the study began. Pretesting was conducted in the language the child was most comfortable with and number knowledge pretest scores, as well as classroom placement, were used to establish three matched groups. One group received the Rightstart program. Control group 1 received an equal amount of small-group attention with a more traditional math program that was specifically designed to provide a level of affective engagement that was commensurate with the Rightstart program. Control group 2 received a language program designed with similar criteria in mind and administered in a similar format.

Studies 2 and 4 were conducted in three public inner-city schools in central Massachusetts that drew children from a variety of cultural backgrounds (e.g., Hispanic, Southeast Asian, Afro- and Anglo-American). All children included in these samples were comfortable with the English language and the number knowledge pretest (administered in all studies in the child’s preferred language) was used, along with the other criteria mentioned above, to establish matched treatment and control groups. In these studies, the control groups were given no additional training beyond their regular classroom instruction. In all studies, matching was nearly perfect for number knowledge performance and cultural background, and less perfect with respect to age. With the exception of study 4, where the Rightstart program was administered in a whole-class format and the control group was drawn
Assessment Objectives and Procedures
Our program evaluation was designed with three objectives in mind. First, we wanted to see whether the RightStart program was sufficient to teach the knowledge specified in the central conceptual structure and whether it did so more efficiently than other, more traditional, kindergarten mathematics programs. Second, we wanted to see whether the knowledge acquired was central to children's performance on a broad range of tasks with some quantitative component, as suggested by the theory. Third, we wanted to see whether the knowledge acquired would enhance children's ability to profit from early formal instruction in arithmetic, as the theory also predicted.

To realize these objectives, we administered a battery of tasks to all children in each sample. The number knowledge test was administered before and after training and the pre-posttest gains were used as a measure of the extent to which children had acquired the knowledge implied in the central conceptual structure. In two studies, children's responses to one item on this test were further analyzed to obtain a measure of strategy use and a probe (i.e., "How did you figure that out?") was included in the test administration to obtain additional information for this analysis.

To see whether the knowledge acquired was central to children's performance in content domains that had not been addressed in the RightStart program, we administered a battery of developmental tasks (i.e., those mentioned in a previous section) after training in all studies and before training in the majority of studies. These included the balance beam task, which requires predictions of relative weight; the birthday party task, which requires predictions of relative happiness; the time knowledge task, which poses a variety of time problems; the money knowledge task, which poses a variety of money problems; and the distributive justice task, which requires attention to issues of fairness.

Each of these tasks has some quantitative requirement in that a successful solution to the problem posed (e.g., "Which side of the beam will go down?" "Who will be happier?" "What time is shown on this clock?") is possible only when the problem array is quantified in some fashion. Each of these tasks also has specific content that was purposely excluded from the RightStart program. In the entire curriculum, any mention of balance systems, birthday parties, time, and money was scrupulously avoided. Our purpose in administering these tests was to see whether the central conceptual structure—if it could
be taught—had the broad range of application the theory suggested. These tasks are described in greater detail in Griffin, Case, and Capodilupo (in press).

The third question we sought to answer was, Does this knowledge enable children to profit from early formal instruction in arithmetic? To answer this question, we retested one sample of treatment and control children (i.e., those included in study 2) one year later, at the end of grade one. Because many children had moved out of the geographic area, the sample was appreciably reduced when the first grade follow-up study was conducted. The children who were located (11 from the treatment group and 12 from the control group) were attending six different schools within a 30-mile radius, and were in 12 different classrooms. The children's first grade teachers had no knowledge of the Rightstart project, or of the treatment status of the children.

To assess arithmetic knowledge at the end of grade one, four instruments were developed or adapted. These included: an oral arithmetic test (e.g., "How much is 2 + 4?"); a written arithmetic test, which presented simple arithmetic problems in a typical worksheet fashion; a word problems test, which was a modified version of one developed by Riley and Greeno (1988) and which required children to provide verbal answers to orally presented word problems; and a teacher rating scale, which required the classroom teacher to rate each child's mathematics performance, in relation to other children in the class, on eight items taken from the school report card. The number knowledge test was also included in this battery to assess the stability of the knowledge acquired in kindergarten, over a one-year period.

Results and Discussion

Number Knowledge

Table 23 shows the percentage of children passing the number knowledge test, before and after training, in five separate studies conducted over a three-year period. As the table indicates, almost all children included in each sample failed the number test prior to training. Four or five months later, the vast majority of the children who had received the training passed the test. By contrast, only a minority of the children in the control groups passed. Note that two of the control groups (i.e., the groups established in study 1) had received training programs that provided an equal amount of individual attention and that contained the same number of games as those developed for the Rightstart program. Multivariate analyses indicated that the gains made by the
<table>
<thead>
<tr>
<th>Study</th>
<th>(n)</th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study 1</td>
<td>20</td>
<td>15</td>
<td>80</td>
</tr>
<tr>
<td>Study 2</td>
<td>23</td>
<td>0</td>
<td>87</td>
</tr>
<tr>
<td>Study 3</td>
<td>7</td>
<td>0</td>
<td>71</td>
</tr>
<tr>
<td>Study 4</td>
<td>38</td>
<td>7</td>
<td>53</td>
</tr>
<tr>
<td>Study 5</td>
<td>10</td>
<td>10</td>
<td>70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study</th>
<th>(n)</th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study 1</td>
<td>20</td>
<td>15</td>
<td>37</td>
</tr>
<tr>
<td>Language control</td>
<td>20</td>
<td>15</td>
<td>35</td>
</tr>
<tr>
<td>Study 2</td>
<td>24</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>Study 4</td>
<td>38</td>
<td>7</td>
<td>14</td>
</tr>
</tbody>
</table>

* Groups drawn from twelve classrooms in six schools located in Canada, California, and Massachusetts.

treatment groups were superior to the gains made by the control groups at the .05 level or better, in each study.

The consistency of these findings, across five separate studies, indicates that the curriculum did what it was supposed to do with considerable effectiveness. Even in the fourth study—where training was conducted with a ratio of one teacher to twenty-five children—the gains were substantial, although lower than the gains shown in the other studies where training was conducted in small groups. Since the curriculum was taught by several different teacher-trainers, to groups of children from widely divergent cultural backgrounds (e.g., Portuguese immigrants in Toronto, Caribbean immigrants in Massachusetts, Afro-Americans in California, and Anglo-Canadians or -Americans in all sites), these findings indicate that the curriculum enabled a wide variety of children to construct, and to consolidate, the knowledge implied in the central conceptual structure. It also appeared to be effective in a variety of educational contexts.

**Strategy Use**

Strategy use was examined on the number knowledge test item (i.e., “If you had 4 chocolates and I gave you 3 more, how many would you have altogether?”) that assessed arithmetic knowledge most directly and that assessed children’s ability to apply the +1 increment rule they
had been taught in the Rightstart program to problems with larger addends, it is worth noting here that none of the activities in the Rightstart program provided training in incrementing quantities larger than one, and the activities that addressed this knowledge objective most directly were the games, which required a small-group instructional format.

Two indices were adopted to obtain a global measure of "reasonable" versus "wild guess" strategy use. The major index was children's responses to the 4 + 3 test item itself. For this measure, a response of 5, 6, 7, or 8 was coded as reasonable on the assumption that each was in the neighborhood of the correct answer and each indicated some awareness that the largest addend should be incremented. A response of "I don't know" was also coded as reasonable on the assumption that it indicated some awareness that the problem required a particular strategy that was not yet available to the child. All other responses, which ranged from 0 to 400, were coded as wild guesses.

The second index available in two studies was children's response to the probe "How did you figure that out?", which was included in the posttest administration of this item. As might have been expected, this question was difficult for many children to answer and these data were used only to provide supporting evidence for the coding categories described above. Children's responses to the probe included "I just knew," "I guessed," "I counted," "I don't know," and blank stares. An analysis of these results indicated that the majority of children who produced responses that were coded as wild guesses also responded to the probe with an "I just guessed" response or a blank stare, lending convergent support to the coding categories that were established for this analysis.

Table 2.4 shows the percentage of children producing wild guess responses to this item in study 2 (small-group instructional format) and study 4 (whole-class instructional format). These data indicate a substantial difference between the treatment and control groups in both studies and a difference of greater magnitude in study 2, where the program was taught in its entirety and in a more ideal instructional

<table>
<thead>
<tr>
<th>Treatment group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study 2 (n = 47)</td>
<td>Study 4 (n = 72)</td>
</tr>
<tr>
<td>4 (87)</td>
<td>11 (53)</td>
</tr>
<tr>
<td>35 (21)</td>
<td>38 (25)</td>
</tr>
</tbody>
</table>

Table 2.4
Percentage of children using "wild guess" strategies on number knowledge test item (numbers in brackets indicate percent children achieving correct answer of 7)
format. Although pretest measures on strategy use were not available, the groups were well matched on several other indices of mathematical knowledge, and it seems reasonable to assume that pretest strategy use was comparable as well.

These data suggest that the children who received the Rightstart program were better able than children who had not, to employ a reasonable strategy in their attempts to solve a novel addition problem and to achieve a higher success rate on this problem. Since children who received the program also gave evidence they had acquired the central conceptual structure the program was designed to teach, it seems reasonable to attribute the superior strategy use of these children to a conceptual knowledge base that enabled them to make sense of the problem and to adopt, thereby, a reasonable strategy to solve it.

Transfer Effects
Table 2.5 shows the percentage of children passing five transfer tests before and after training. The data used for this table were an aggregate of three sets of findings, collected over a three-year period, and they provide a reasonable estimate of transfer effects in each study. As the table indicates, very few children were able to pass any of these tests prior to training. By contrast, on the posttest, the majority of the treatment group passed four of the five transfer tests and the majority of the control group failed. Even on the money knowledge test—where the absolute magnitude of children passing was lower—the difference between the treatment and control groups was still significant at the .01 level. Multivariate analyses provided similar or higher confidence levels on each of the remaining tests. These findings provide strong evidence that the knowledge implied in the central conceptual structure, and acquired through our readiness training, has the broad range of application the theory suggests.

Table 2.5
Percentage of treatment and control children passing five transfer tests (aggregate data from five studies)

<table>
<thead>
<tr>
<th></th>
<th>Pretest</th>
<th>Posttest treatment group</th>
<th>Posttest control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balance beam</td>
<td>(n = 183)</td>
<td>10</td>
<td>77</td>
</tr>
<tr>
<td>Birthday party</td>
<td>(n = 107)</td>
<td>16</td>
<td>81</td>
</tr>
<tr>
<td>Tune-telling</td>
<td>(n = 183)</td>
<td>11</td>
<td>65</td>
</tr>
<tr>
<td>Money-handling</td>
<td>(n = 183)</td>
<td>8</td>
<td>38</td>
</tr>
<tr>
<td>Distrib. justice</td>
<td>(n = 47)</td>
<td>—</td>
<td>87</td>
</tr>
</tbody>
</table>
First Grade Follow-Up

When the number knowledge test was readministered, one year later, to the first grade follow-up sample, the results indicated that all of the children who had received the Rightstart program in kindergarten passed level 1 of this test at the end of grade one (versus 87% passing one year earlier). At this point in time, 83 percent of the control group also demonstrated this knowledge (compared to 25% one year earlier). These findings suggest that: first, the knowledge the treatment children acquired in kindergarten was stable over a one-year period and was very likely available to them when they started first grade; and second, at some point during their first grade experience, many children in the control group also acquired this knowledge. These findings raise an interesting question, namely: If many children can acquire the knowledge taught in the Rightstart program in the course of their first grade experience, why bother to teach it in kindergarten? The results described below provide an answer to this question.

Table 2.6 presents the results of the number knowledge test and the arithmetic tests that were included in the first grade battery. These findings indicate that some children in the treatment group were now able to solve double-digit problems that were included in level 2 of the number knowledge test. No child in either group passed these items at the end of kindergarten and none of the control children passed them at the end of grade one.

When children’s performance on the first grade arithmetic tests is examined, the findings indicate that the majority of the treatment group passed two of these tests, the oral arithmetic test and the word problems test, whereas a large proportion of the control group failed.

<table>
<thead>
<tr>
<th></th>
<th>Treatment group (N = 23)</th>
<th>Control group (N = 54)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number knowledge:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level 1</td>
<td>100</td>
<td>83</td>
</tr>
<tr>
<td>Level 2</td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td><strong>Oral arithmetic</strong></td>
<td>82</td>
<td>33</td>
</tr>
<tr>
<td><strong>Written arithmetic</strong></td>
<td>91</td>
<td>75</td>
</tr>
<tr>
<td><strong>Word problems:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level 1</td>
<td>96</td>
<td>54</td>
</tr>
<tr>
<td>Level 2</td>
<td>46</td>
<td>13</td>
</tr>
</tbody>
</table>
Table 2.7
Percentage of treatment and control children achieving teacher ratings of average or above at the end of grade one

<table>
<thead>
<tr>
<th></th>
<th>Treatment group (n = 11)</th>
<th>Control group (n = 12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number sense</td>
<td>200</td>
<td>24</td>
</tr>
<tr>
<td>Meaning of numbers</td>
<td>88</td>
<td>42</td>
</tr>
<tr>
<td>Use of numbers</td>
<td>88</td>
<td>42</td>
</tr>
<tr>
<td>Addition</td>
<td>100</td>
<td>66</td>
</tr>
<tr>
<td>Subtraction</td>
<td>100</td>
<td>66</td>
</tr>
<tr>
<td>Accuracy</td>
<td>63</td>
<td>59</td>
</tr>
<tr>
<td>Speed</td>
<td>75</td>
<td>42</td>
</tr>
</tbody>
</table>

On each test, the groups differed significantly. On the third test, the written arithmetic test, a large proportion of both groups passed and the difference between the groups was not significant. Since the written test required children to solve the sort of worksheet problem they encountered on a daily basis in their classrooms, this finding is not surprising. It suggests that most children were able to profit from extensive exposure to a particular problem format.

On the teacher rating scale, almost all children in the treatment group received a rating of average or above average on five of the seven items included on this scale (see table 2.7). The proportion of children in the control group receiving this rating was significantly lower. The differences between the groups are especially striking on the three items that indexed the conceptual knowledge the Rightstart program was designed to teach, namely: “demonstrates number sense,” “understands the meaning of numbers,” “understands the use of numbers.” Conversely, on the two items that were not addressed or considered important in the Rightstart program—“works with reasonable speed,” “works accurately”—there was no difference between the groups. The high ratings the treatment children received from their first grade classroom teachers are also striking when it is recalled that these children had been performing in the below-average category in kindergarten and had been included in the Rightstart training sample for this reason.

When the data from the first graders are considered as a whole, they suggest that both groups (i.e., treatment and control) acquired some new knowledge in their first grade experience. Both groups learned to solve the sort of arithmetic problem they encountered on a daily basis in their classroom. However, when children were required to use this knowledge more flexibly, to solve formal problems presented orally, to
solve word problems, or to demonstrate to their classroom teachers that they possessed number sense, a large proportion of the control group was unable to do so. By contrast, the treatment children appeared to have acquired new knowledge that was general, and that could be applied in a more flexible fashion.

One possible interpretation of these findings would be to suggest that the treatment children had a central conceptual structure in place, at the beginning of grade one, which lent meaning to the facts, algorithms, and strategies that were taught in their first grade classrooms. By assimilating this procedural knowledge into their conceptual structure, they were able to construct new knowledge that was general, and that could be deployed in a variety of situations. Lacking such a structure, the control children were able to master some of the procedural knowledge that was taught in their first grade classrooms, but were unable to make sufficient sense of it to permit generalization to new problem contexts.

We ourselves were surprised at the magnitude of these findings. We had hoped that the Rightstart program would have some long-term effects, but we had also feared that these effects might be lost when children were exposed to a whole year of traditional instruction that provided little support for the conceptual understandings taught in the kindergarten year. The findings provide evidence that training effects are robust and that the Rightstart program does, indeed, facilitate first formal learning of arithmetic.

Conclusions

Taking the data we have gathered across the past three years as a whole, we draw the following general conclusions:

1. Children from different social strata in America enter school with substantial differences in their understanding of numbers. Those in middle and high SES groups view the number system in the dimensional fashion indicated in the central conceptual structure (figure 2.2). Many children from low-income families appear to lack this conceptual knowledge; they see numbers in a predimensional fashion or they cannot apply their dimensional understanding in the sort of context that our schools provide. Given the problems that American schools have in fostering conceptual understanding in children from all backgrounds (Stigler, Lee, and Stevenson 1990), this sort of early difficulty seems particularly unfortunate, since it increases the likelihood that children will treat school math from the start as a rote activity.
2. The Rightstart program appears to be effective in eliminating these differences and in enabling a wide variety of children, who come to school from different starting points and/or with different linguistic conventions, to acquire the conceptual knowledge it was designed to teach. Its effectiveness has now been demonstrated in two countries (Canada and the United States), in six separate schools, and with five lower- and lower-middle-income groups. It has also been shown to be effective when implemented by eight different teacher-trainers.

3. It seems clear that the program does not have its effect simply by familiarizing children with the procedure for counting, or for representing numbers with numerals. Nor does it have its effect simply by engaging children in a social process with which they have had little experience up to that point in time, namely, answering test questions by an adult, and justifying these answers in a “decontextualized” setting. If these were the only important aspects of the program, then one would have expected more progress to be shown by the control groups that were used in study 1.

4. Three of the main foci of the program were (1) teaching children to respond to questions about relative magnitude in the absence of any concrete sets of objects, (2) teaching children the “increment rule,” i.e., the rule that dictates that the addition or subtraction of one element to a set alters the cardinal value of that set by one unit, and therefore moves the value one unit up or down on the number line, and (3) teaching children that knowledge of relative position on the number line is useful for determining relative quantity in various “real world” tasks, when it cannot be determined more directly. Since the sorts of questions on which children showed improvement on the post-test battery all dealt with one or another of these components, it can be assumed that the program had its effect for the reason hypothesized; namely, it enabled children to acquire the numerical understandings specified in the central conceptual structure.

5. On the basis of the transfer effects, it seems clear that the knowledge taught is central to children’s performance on a range of quantitative tasks for which no specific training had been provided.

6. On the basis of the first graders’ performance, it also seems clear that positive effects of the Rightstart program are apparent one full year after the program has ended. After being exposed to a standard first grade arithmetic curriculum in a variety of classrooms in a variety of schools, children who had received the
program in kindergarten did significantly better than their peers in the control group on several achievement measures administered at the close of first grade. The mathematics knowledge and learning they demonstrated at this point was commensurate with that of their middle-income peers and in sharp contrast to the learning potential these children had demonstrated in the middle of their kindergarten year.

In summary, the major conclusion we draw from these findings is a straightforward one. Although children from different backgrounds may still need different sorts of programs to accommodate their needs as they progress through formal schooling, the Rightstart program ensures that children from a diverse array of backgrounds will all start first grade with an understanding of quantities, numbers, and numerical terminology that builds on their existing insights and vocabulary, and that is well matched to the requirements of first grade.

Note

The Rightstart program and the research that supported its development and evaluation were made possible by a grant from the James S. McDonnell Foundation.