A Microgenetic/Cross-Sectional Study of Matrix Completion: Comparing Short-Term and Long-Term Change

Robert S. Siegler and Matija Svetina

A design that included both microgenetic and cross-sectional components was used to examine 135 Slovenian children’s acquisition of matrix completion proficiency and compare microgenetic and age-related changes on the task. The microgenetic analyses indicated that children’s errors became increasingly variable shortly before they discovered the correct strategy, that the correct strategy became dominant quite quickly following its initial use, that improvements in matrix completion performance generalized to conservation, and that amount of learning correlated positively with IQ. The microgenetic/cross-sectional comparison, which involved contrasting the changes that occurred over seven experimental sessions with the changes that occurred between ages 6 and 7 years, indicated that the two patterns of change were highly similar.

INTRODUCTION

This study had two main purposes. The first was to enhance understanding of acquisition of matrix completion proficiency through the use of microgenetic analysis. The second was to compare the changes in these capabilities elicited by the microgenetic intervention to the changes that occur with age in the absence of such an intervention.

The Microgenetic Method

Over the past 2 decades, an increasing number of investigators have adopted the microgenetic approach as a means for studying cognitive development (for a review, see Miller & Coyle, 1999). The main reason is the precise descriptions of changing competence that the approach can yield. The microgenetic method is defined by three primary characteristics: Observations span the period of rapidly changing competence; the density of observations within this period is high, relative to the rate of change; and observations of changing performance are analyzed intensively to indicate the processes that gave rise to them.

The second characteristic is especially important. A dense sampling of performance while the performance is changing provides the high temporal resolution needed to describe the process of change. The detailed data about changes yielded by microgenetic analyses also allows investigators to discriminate between alternative underlying mechanisms. Many mechanisms could potentially produce changes in a single index of cognitive growth, such as percentage of correct answers. Far fewer mechanisms, however, could give rise to the highly specific data about changes in strategy use, particular errors, solution times, and range of generalization that can emerge from such studies.

Microgenetic methods have proven to be applicable for studying development in diverse content areas: language (Gershkoff-Stowe & Smith, 1997), memory (Coyle & Bjorklund, 1997), attention (Miller & Aloise-Young, 1996), locomotion (Thelen & Ulrich, 1991), arithmetic (Siegler & Jenkins, 1989), and scientific reasoning (Kuhn & Phelps, 1982) among them. They also have been shown to be applicable to people of a wide age range: infants (Adolph, 1997), toddlers (Chen & Siegler, 2000), preschoolers (Johnson & Mervis, 1994), school-age children (Schauble, 1996), adolescents (Kuhn, Garcia-Mila, Zohar, & Andersen, 1995), and adults (Granott, 1998).

In addition to this broad applicability, microgenetic studies have suggested a general conceptual framework for thinking about cognitive development (Siegler, 1996). This framework distinguishes among five dimensions of cognitive growth: its path, rate, breadth, source, and variability. The path of change concerns the sequence of knowledge states or problem-solving approaches that children use while gaining competence. The rate of change involves the period of time or amount of experience that separates initial use of a new approach from consistent use of it. The breadth of change involves how widely the new approach is generalized to other problems and contexts. The source of change concerns the causes that set the change in motion. The variability of change involves differences among children in the other dimensions of change.

An especially encouraging characteristic of micro-
genetic studies is that despite the diversity of content areas and populations associated with their use, they have yielded surprisingly consistent findings (for reviews, see Kuhn, 1995; Miller & Coyle, 1999; Siegler, 2000). A common finding regarding the path of change is that just before discovery of a new approach, performance becomes more variable (Alibali, 1999; Goldin-Meadow, Alibali, & Church, 1993; Graham & Perry, 1993; Siegler, 1995). The rate of change tends to be gradual, with less sophisticated, earlier emerging approaches continuing to be used well after more sophisticated approaches also have emerged (Bjorklund, Miller, Coyle, & Slawinski, 1997; Kuhn et al., 1995; Siegler & Jenkins, 1989). The breadth of change usually is fairly narrow (Kuhn et al., 1995; Schauble, 1990, 1996; Siegler & Jenkins, 1989). Variability tends to be high: children learn via different paths, at different rates, and with differing amounts of generalization. Finally, certain sources of cognitive growth, such as encouragement to explain observations of physical phenomena or other people’s reasoning, operate over a wide age range and in diverse content domains (Bielaczyc, Pirolli, & Brown, 1995; Chi, de Leeuw, Chiu, & LaVancher, 1994; Pine & Messer, 2000; Siegler, 1995, in press). Thus, examining cognitive growth along the five dimensions has proven to be useful for identifying regularities in how change occurs.

Development of Matrix Completion

The present research employed microgenetic methods to study matrix completion, a logical reasoning task prominent within Piagetian, information processing, and psychometric traditions alike (Carpenter, Just, & Shell, 1990; Inhelder & Piaget, 1964; Raven, Court, & Raven, 1995). The matrix completion task used in the present study was modeled after one used by Inhelder and Piaget. As illustrated on the left side of Figure 1, each problem involved a $2 \times 2$ matrix, with objects occupying the top row and the left column; the square in the lower right quadrant of the matrix was empty. Six alternatives that could potentially complete the matrix (right side of Figure 1) also were presented. The task was to choose the answer that, if inserted into the empty space in the $2 \times 2$ matrix on the left, would create the same relations between the two objects on the bottom, as already exists between the two objects on the top (and between the two objects on the right as already exists between the two objects on the left). The correct answer is the small, light-gray bird facing left.

![Figure 1: Example of a matrix completion problem.](image)

Figure 1 Example of a matrix completion problem. The task is to choose an object from the six potential answers on the right that, if inserted into the blank space in the $2 \times 2$ matrix on the left, would create the same relations between the two objects on the bottom, as already exists between the two objects on the top (and between the two objects on the right as already exists between the two objects on the left). The correct answer is the small, light-gray bird facing left.

to a large, light-gray bird facing right, as a small, light-gray mouse facing left is to a ______. A child who generated this description could then search for a small, light-gray bird facing left and select the alternative that matched that description (the bird at the bottom of the left column of potential answers in Figure 1).

The development of ability to solve matrix completion problems is of interest for several reasons. From the perspective of Piagetian theory, the task measures in a particularly direct way the ability to focus on multiple dimensions rather than on just a single one. Consistently correct performance requires children to decenter and consider all dimensions, less salient as well as more salient. From the perspective of information-processing theory, successful matrix completion requires appropriate encoding and analogical reasoning, two pervasively important processes in learning and problem solving (e.g., Gentner, 1989; Halford, 1993; Sternberg, 1977). From the perspective of psychometric theory, matrix completion performance, as measured by the Raven’s Coloured Progressive Matrices Test (CPMT; Raven, Court, & Raven, 1995), correlates highly with scores on a wide variety of ability tests, as well as on subtests of specific skills such as ability to draw verbal, geometric, and numerical analogies and to induce letter series (Jensen, 1987; Marshalek, Lohman, & Snow, 1983; Snow, Kyllonen, & Marshalek, 1984). Thus, matrix completion seemed a worthwhile task to study.

The development of matrix completion competence was first examined intensively by Inhelder and Piaget (1964). They found that substantial develop-
ment occurs between the ages of 5 and 8 years. The path of change that they described was not monotonic; 6-year-olds did fairly well, 7-year-olds did less well, and 8-year-olds did best of all. Inhelder and Piaget also noted that before children began to respond correctly, they often chose duplicates of the object directly above or alongside the blank square in the matrix (e.g., the mouse in the middle of the right column of potential answers in Figure 1 is a duplicate of the mouse on the bottom left of the matrix). A third observation was that choosing objects that were properly oriented posed greater difficulty than choosing ones of the correct form, size, and color.

Inhelder and Piaget’s (1964) findings triggered a flurry of research on matrix completion over the next 15 years. These studies consistently replicated the initial findings that 5- and 6-year-olds rarely solved such problems accurately and that 8-year-olds were much more accurate (Overton & Brodzinsky, 1972; Parker & Day, 1971). Also replicating Inhelder and Piaget’s findings were results indicating that items that required attention to less salient dimensions, such as orientation, were more difficult than items that only required attention to salient dimensions, such as form and size (Odom, Astor, & Cunningham, 1975). On the other hand, subsequent researchers (e.g., Kingma, 1983) have not replicated the decline in performance between ages 6 and 7 years reported by Inhelder and Piaget.

Subsequent research also extended Inhelder and Piaget’s findings in several directions. Matrix completion performance was found to correlate positively with performance on other measures of logical reasoning, such as conservation (Carlson & Weidel, 1977; Dimitrovsky & Almy, 1975), seriation (Hamel & van der Veer, 1972), and the CPMT (Carlson & Goldman, 1974). In addition, training studies indicated that presenting verbal rules for solving matrix completion problems and feedback concerning applications of the rules enabled children as young as 6 years to succeed on the task (Parker, Sperr, & Rieff, 1972).

These findings provide a broad outline of development of matrix completion. They are less informative about the process of acquisition, however; in particular, its path, rate, breadth, variability, and source. With regard to the path of change, Inhelder and Piaget (1964) identified one important phenomenon—that duplicate errors are prevalent prior to acquisition of matrix completion competence—but subsequent studies have added little to that initial finding. With regard to the breadth of change, subsequent studies have indicated that success at matrix completion is positively correlated with success on other logical reasoning tasks, but leave unclear whether the correlations reflect any causal relation. With regard to the source of change, training studies have demonstrated that highly directive instruction, such as telling children verbal rules for solving such problems, can lead to learning, but such studies tell us little about the effectiveness of less directive approaches that seem more likely to occur in the everyday environment. With regard to the rate and variability of change, almost nothing is known.

To increase understanding of the path, rate, breadth, variability, and sources of acquisition of matrix completion proficiency, 6-year-olds in the present study were presented with either a seven-session microgenetic manipulation or a control condition. As shown in Table 1, in Session 1, children in both experimental and control groups were presented a no-feedback examination of matrix completion competence and also tests of conservation knowledge and IQ. In Sessions 2 through 5, children in the experimental condition received potentially instructive experience with the task (described in the following paragraph). In Session 6, these children were tested under the same no-feedback conditions as in Session 1; the purpose was to measure their performance immediately after the instructional experience. In Session 7, 2 months later, matrix completion and conservation problems were presented to children in both the ex-

<table>
<thead>
<tr>
<th>Session</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-year-olds</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: MC = matrix completion; C = conservation; R = Ravens Progressive Matrices Test.

a Feedback and self-explanation questions given.
Comparing Microgenetic and Age-Related Change

The second major purpose of the current study was to present and illustrate the usefulness of a method for comparing microgenetic and age-related changes. There is widespread agreement that the short-term changes that emerge in microgenetic studies resemble to some degree the longer-term changes that emerge in longitudinal and cross-sectional studies (e.g., Fischer & Biddell, 1998; Granott, 1998; Miller & Coyle, 1999). There is much less agreement, however, as to the level of detail at which the resemblance holds. Miller and Coyle (Pressley, 1992), and Kuhn (1995) all have noted that the degree of similarity between microgenetic and age-related change is uncertain, both descriptively and at the level of underlying mechanisms. These issues led Miller and Coyle (p. 212) to conclude: “Although the microgenetic method reveals how behavior can change, it is less clear whether behavior typically does change in this way in the natural environment” (italics in original).

After noting this uncertainty with regard to the extent of resemblance between microgenetic and age-related change, Fischer and Granott (1995) and Kuhn (1995) suggested a reasonable solution: Compare the particular changes observed in microgenetic studies to those observed in cross-sectional and long-term longitudinal studies. As they noted, such comparisons indicate considerable similarity: the same strategies typically emerge; the breadth of generalization tends to be fairly narrow; and substantial variability is present, both between and within individuals. Differences across studies in populations, methods, and measures, however, have limited the level of detail at which the short- and long-term changes could be compared.

Kuhn’s (1995) and Fischer and Granot’s (1995) argument can be extended in a way that would allow more precise comparisons between the two types of change than has previously been possible. The basic strategy is to combine cross-sectional and microgenetic components within a single experimental design, using the same population, tasks, instructions, and measures of performance. Table 1 illustrates the design used to implement this logic in the present study. In the first session, 6-, 7-, and 8-year-olds were given the task of primary interest (the matrix completion task), as well as tests of conservation knowledge and intelligence. This session provided all of the data for the cross-sectional component of the study, as well as the pretest measures for the microgenetic component. Sessions 2 through 7 completed the microgenetic portion of the study, as described earlier. Be-
cause children in the cross-sectional and microgenetic portions of the study were chosen from the same population, and because identical tasks, instructions, and measures were used in both, it was possible to evaluate, in an unusually direct way, the similarity between microgenetic and age-related change.

The combination of microgenetic and cross-sectional components enabled us to directly compare changes on 10 measures of matrix completion performance: percentage of correct answers; percentage of answers that were correct on the dimensions of form, size, orientation, and color; percentage of explanations that cited each of these four dimensions; and predominant type of errors. The data also allowed for two additional comparisons of microgenetic change to the pattern of age-related change predicted by Piagetian theory. One comparison involved stability of learning over time; the other involved generalization of learning to novel tasks. Piaget (e.g., 1964) identified stability and generalization as key properties of development (as opposed to learning, which would not be expected to show either property). Thus, a total of 12 measures were examined in this comparison. The present design allowed us to examine whether learning within microgenetic studies also showed these characteristics.

METHOD

Participants

The children in the present study were 6-, 7-, and 8-year-olds, who were recruited from kindergartens and primary schools in Ljubljana, Slovenia. The sample originally included 157 children: ninety 6-year-olds (M = 72.3 months, range = 70–74), thirty-seven 7-year-olds (M = 84.5 months, range = 82–86), and thirty 8-year-olds (M = 96.0 months, range = 94–98). Among the ninety 6-year-olds, ten did not understand the example problems (described below) and did not participate further in the study. The remaining eighty 6-year-olds were randomly assigned to the control or experimental group. Two of the forty 6-year-olds originally assigned to the control group missed one of their two sessions, however, and 10 of the 40 children assigned to the experimental group missed at least one of their seven sessions. Thus, the analyses of the 6-year-olds’ performance were based on the data from the 38 children in the control group and the 30 in the experimental group who were present at all sessions.

The four groups of children (6-year-olds in the experimental and control groups and 7- and 8-year-olds who received Session 1 only) were comparable with regard to parental educational level (as indicated by parents’ self-reports). On a scale in which completion of elementary school was scored as 1, high school as 2, and college as 3, mean parental education was 2.36. A 30-year-old male graduate student conducted the experiment.

Tasks and Procedure

Presentation of tasks followed the plan described in Table 1. Sessions 1 through 6 were presented at 3- to 5-day intervals. Session 7 was presented 7 to 8 weeks after Session 6 (10 weeks after Session 1). Session 1 was presented over 2 successive days; all subsequent sessions were presented on a single day.

Matrix completion. To ensure that children understood the matrix completion task, eight example items were presented at the beginning of Session 1. Each example involved a $2 \times 2$ matrix with objects in three of the four squares; the bottom right square was empty. The items in a given example matrix varied along only one dimension, and the experimenter described each matrix in a way that conveyed this fact. For example, one introductory matrix included a large green fish facing left in each of the two squares on the left and a small green fish facing left in the top right square. The experimenter said, “Look: This is a big fish and this is a little fish [pointing to the top left and top right squares]. Here there is also a big fish [pointing to the bottom left square], and here there is one missing [pointing to the bottom right square]. What do you think this one should look like [pointing to the empty square]?” After the child answered, the experimenter asked, “OK, which of these do you think is correct,” while pointing to the six alternative answers, which were arranged in three rows and two columns. The 10 children who did not correctly answer at least half of these simple items were assumed to not understand the task and were not tested further.

Children who met the criterion for understanding the task were then presented 22 test matrices in each session in which they participated. (Sessions 1–7 for children in the experimental group, Sessions 1 and 7 for children in the control group.) These items varied in two, three, or four dimensions. A matrix’s number of dimensions of variation reflected the ways in which the three objects in the matrix differed from each other. To illustrate, the Figure 1 matrix varies in three dimensions: form (mouse or bird), size (large or small), and orientation (facing left or right). Whether a matrix varied in two, three, or four dimensions proved to be unrelated to children’s success in solving it; therefore, this variable is not discussed further.

Before the first test problem, the experimenter said,
“There are three animals here [pointing to the objects in the matrix]. Here [pointing to the empty cell], one is missing—one of these [pointing to the six alternative answers]. Which one is missing?” Once the child answered, the experimenter asked, “Why do you think so?” On subsequent problems, the instructions were shortened, with the experimenter simply asking, “Which of these [pointing to the potential answers] belongs here?” The six alternative answers for each item included the correct answer (the alternative that had the right value on all four dimensions), four answers that had the correct value on three of the four dimensions but not on the fourth, and one answer that was incorrect on all four dimensions. Therefore, although only one of the six choices was correct on all four dimensions, four of the six choices were correct on any single dimension (e.g., four of the six potential answers had the correct orientation). These response alternatives made it possible to examine whether children’s incorrect choices stemmed from their paying insufficient attention to a given dimension. For example, if children picked the answer that was wrong only on the orientation dimension more often than other incorrect answers, we could infer that they were attending less to orientation than to the other dimensions.

The variation among the six potential answers to each problem should not be confused with the variation among the three items within the $2 \times 2$ matrix. The six potential answers always varied on all four dimensions (size, form, orientation, and color), whereas the three items within the matrix could vary on two, three, or all four dimensions. For example, in Figure 1, the three items within the matrix were all gray, so they did not vary in color; they varied only in the three dimensions of size, form, and orientation.

Children were scored as citing a dimension in an explanation if they mentioned it in their initial statement. Only explanations of incorrect answers were examined, because children who responded consistently correctly often became bored with citing all of the dimensions and began advancing explanations such as “because it’s the right one” trial after trial. Efforts to motivate them to elaborate on these cryptic explanations generally were unsuccessful.

Intelligence test. The booklet form of the CPMT (Raven et al., 1995) was administered in Session 1 to determine whether children’s learning of the present matrix completion task was related to their performance on a well-standardized psychometric instrument with similar content. For U.S. first and third graders, the CPMT has a split-half reliability of .85, and a test–retest reliability ranging from .81 to .87 over periods ranging from 10 days to 1 month.

The fact that this intelligence test, like the main experimental task, included matrix completion items offered both advantages and disadvantages. The main advantage was that the similar formats maximized the likelihood of detecting a relation between performance on the experimental task and on a broadly predictive standardized test. The main disadvantage was that the similar formats could have led to this relation being stronger than it would have been if a different standardized test had been used. Given the limited current knowledge regarding the relation between psychometric test scores and microgenetic performance, the costs involved in this trade-off seemed acceptable.

Number conservation. In Sessions 1 and 7, children were presented with two conservation problems, one on number conservation and the other on liquid quantity conservation. The number conservation problem closely resembled one described in Piaget (1952). Children were shown two parallel rows, each containing seven coins. The coins in the two rows initially were in 1:1 correspondence and were arranged horizontally in front of the child. The experimenter said “Here we have two rows of coins. Are there the same number of coins in this row as in this one?” When the child agreed that the number of coins was equal, the experimenter spread out the row of coins closer to the child and asked, “Are there the same number of coins now?” After the child answered, the experimenter asked, “How do you know?” The experimenter then said “OK,” regardless of the child’s response. Answers were considered correct if children said that the number of coins was the same and explained their reasoning in terms of the reversibility of the operations, the lack of addition or subtraction, or the identity of the objects before and after the transformation.

Liquid quantity conservation. Children were presented with a liquid quantity problem that closely paralleled another task described in Piaget (1952). The experimenter initially presented three clear glasses. Two were short, wide, and identical to each other; the third was tall and narrow. Initially, the short, wide glasses contained water, and the tall, narrow glass was empty. The experimenter pointed to the short, wide glasses and asked, “Is there the same amount of water in both glasses?” If children said “no,” they were asked to make the amounts equal. Then, the experimenter poured the water from one short, wide glass into the tall, narrow glass and said, “Look, I poured all of the water from this glass into this one. Is there now the same amount of water in both glasses or not?” The child’s answer was followed by the question, “How do you know?” The
experimenter again answered “OK,” regardless of the child’s answer. Criteria for a correct response were the same as with the number conservation problem.

Experimental conditions. The difference between the experimental and control conditions occurred in Sessions 2 through 6, which were presented only to those 6-year-olds in the experimental group (Table 1). During Sessions 2 through 5, the experimenter provided children with feedback following each of their answers and then requested that they explain why the correct answer was correct. In particular, after correct answers, the experimenter said “OK, that is correct” and then asked “Why was it correct?” After incorrect answers, the experimenter said “No, I would choose this one” (pointing to the right answer) and then asked “Why do you think I would pick that one?” Regardless of the child’s explanation, the experimenter responded “OK.”

In Session 6, matrix problems were presented under conditions identical to those under which the problems were presented in Sessions 1 and 7. The purpose was to provide a measure of the amount of learning of children in the experimental condition at the end of the feedback sessions. Results of Session 6 also provided a baseline for examining the stability of learning over the 7 to 8 weeks between Sessions 6 and 7.

RESULTS

The results are presented in three sections. In the first, we describe findings from the microgenetic portion of the study. In the second, we describe findings from the cross-sectional portion. In the third, we compare the microgenetic and age-related changes. Where not otherwise stated, post hoc comparisons are Newman-Keuls tests with $p < .05$.

Microgenetic Changes

Findings from the microgenetic portion that was examined were analyzed in terms of five dimensions of change: source, variability, path, rate, and breadth.

Source of change. The first question that was examined was whether practice, feedback, and self-explanation questions helped children in the experimental group learn how to solve matrix completion problems. To find out, performance of 6-year-olds in the experimental and control conditions in Sessions 1 and 7 was contrasted. A 2 (group: experimental or control) $\times$ 2 (session: 1 or 7) repeated-measures ANOVA on percentage of correct answers on matrix completion problems yielded effects for group, $F(1, 66) = 7.77, p < .01$, session, $F(1, 66) = 17.48, p < .001$, and the Group $\times$ Session interaction, $F(1, 66) = 11.53, p < .001$. As shown in Figure 2, no difference in percentage of correct answers was present on the pretest. In contrast, on the posttest, children in the experimental group answered correctly more often than did children in the control group, $t(66) = 3.60, p < .01$. Viewing the results from a different perspective, children in the experimental group improved from pretest to posttest, $t(29) = 3.70, p < .01$, but children in the control group did not improve during that period. Thus, the experimental manipulation was a source of cognitive change.

Although the experimental manipulation influenced how many errors children made, it did not influence the types of errors they made on those trials on which they were incorrect. Most errors in each session were duplicate errors; that is, responses identical to one of the objects in the matrix. The matrices were constructed so that on 10 of the 22 problems in each session, one of the five possible erroneous responses was a duplicate of an object above or alongside the blank square in the matrix; on the other 12 problems, two of the five erroneous responses were duplicates. This meant that the chance likelihood of an incorrect answer being a duplicate was 31%. Analyses indicated that most errors of children in the experimental group were duplicates (59% versus the 31% that would be generated by chance), $t(27) = 7.22, p < .001$. Duplicates were the predominant error in all sessions, with the percentage of errors that were duplicates not varying greatly from session to session (e.g., 62% in Session 1, 58% in Session 7).

![Figure 2 Percentage of correct answers on the matrix completion task in the first and last sessions among 6-year-olds in the control and experimental groups.](image-url)
Recall that on each problem, four of the six choices were correct on each individual dimension (for example, four choices were the correct size), even though only one choice was correct on all four dimensions. To determine whether the experimental manipulation led to greater improvements on some dimensions than others, a 7 (session) × 4 (dimension) repeated-measures ANOVA on percentage of answers that were correct on each dimension was conducted. (This ANOVA was limited to children in the experimental condition, who were the only ones who participated in all seven sessions.) The analysis revealed a Session × Dimension interaction, $F(18, 522) = 2.09, p < .01$, as well as main effects for session, $F(6, 174) = 3.76, p < .01$, and dimension, $F(3, 87) = 3.35, p < .05$. The interaction was due to the number of correct answers increasing from Session 1 to Session 7 for orientation (72% versus 84%), $F(6, 174) = 6.41, p < .001$, and size (77% versus 84%), $F(6, 174) = 2.82, p < .05$, but not for form (87% versus 88%) or color (83% versus 87%). Seen from another perspective, in Session 1, answers more often were correct on form and color than on orientation and size (87% and 83% versus 72% and 77%, respectively), but by Session 7, answers were correct equally often on all four dimensions (84–88%).

Also examined was the percentage of trials on which children cited each dimension in their explanation of their own response or of the answer that the experimenter said was correct. A Dimension × Session ANOVA revealed main effects for both dimension, $F(3, 39) = 8.54, p < .01$, and session, $F(6, 78) = 2.47, p < .05$. In all sessions, form and color were cited on higher percentages of trials than were size and orientation (71% and 56% of trials versus 72% and 77%, respectively). Number of explanations mentioning each dimension did not change significantly over sessions for any of the four dimensions considered individually.

**Variability of change.** Examination of individual children’s performance over the seven sessions indicated three patterns of change, which were labeled the precocious, learner, and nonlearner patterns. The precocious pattern involved accurate performance from Session 1 onward. The learner pattern involved inaccurate performance in Session 1 but accurate performance by the later sessions. The nonlearner pattern involved inaccurate performance at both beginning and end. All 30 children in the experimental condition fit one of the three patterns of change.

A wide range of criteria would have led to identical classifications of children’s performance within these three patterns. Five children fit the precocious pattern, which was operationally defined as at least 80% of problems correct in the first session and a mean of at least 80% correct answers over the last three sessions. Eight children fit the learner pattern, which was defined as fewer than 33% correct answers in the first session and a mean of at least 80% correct answers over the last three sessions. Finally, 17 children fit the nonlearner pattern, which was defined as fewer than 33% correct answers in the first session and a mean of fewer than 33% correct answers over the last three sessions. Figure 3 indicates the very large differences in performance and learning of the children who fit these three patterns. In Session 1, precocious children answered correctly more often than learners or nonlearners. In Sessions 2 and 3, precocious children answered correctly more often than learners, who answered correctly more often than nonlearners. In Sessions 4 through 7, precocious children and learners answered correctly equally often, and both answered correctly more often than nonlearners.

In addition to these differences on the matrix completion task, children in the three groups differed in their intelligence test scores, $F(2, 27) = 4.66, p < .05$. Precocious students scored in the 84th percentile, learners in the 56th percentile, and nonlearners in the 36th percentile. The difference between precocious children and nonlearners was significant. Results were similar when matrix completion performance was examined as a continuous variable; children’s IQ
scores correlated with their number correct on the matrix completion task, $r(28) = .48$, $p < .05$.

Path of change. Identifying the set of children who learned to solve the matrix completion problems made it possible to analyze their acquisition process at a fine-grain level. In particular, for these “learners”—the one group that showed substantial learning—the path, rate, and breadth of cognitive change were examined.

Examined first was the path of change—the choices (correct, duplicate, or other error) leading up to and following discovery of the correct solution procedure. Each child’s point of discovery was operationally defined as the first set of three successive trials that the child answered correctly. Given that only one of the six alternatives on each problem was correct, the chance probability of generating three consecutive correct answers was $(1/6)^3$ or $1/216$. The particular criterion was unimportant; only one child’s point of discovery would have been classified differently if a criterion of four consecutive correct responses had been used instead. Note that by using this criterion it was possible for a child to discover the correct approach at the very end of Session 1 yet still be classified as a learner, because the child’s percentage of correct answers for the session as a whole was below 33%. This occurred in two cases.

Figure 4 uses a backward trials graphing procedure to illustrate the path of change (Siegler, 1996). The 0 trial block refers to the three trials on which the child first met the criterion for discovery; thus, by definition, performance on that trial block was 100% correct. The –1 trial block refers to the three trials immediately before the 0 block, the –2 trial block refers to the three trials immediately before the –1 block, the +1 trial block refers to the three trials immediately after the 0 block, and so on.

Figure 4 illustrates the overlapping waves pattern typical of strategic change. At first, children usually chose a duplicate of the object above or alongside the blank square. Then, about 12 trials before the discovery, frequency of duplicates dramatically decreased and frequency of other errors increased. Then came the discovery, followed by about 20 problems on which performance was usually correct, although not perfect. After that, performance became consistently correct. To be specific, children made fewer duplicate errors in the four trial blocks immediately before discovery than in the four prior to those, $t(6) = 4.17$, $p < .01$; they made more errors of other types in the four trial blocks immediately before discovery than in the prior four blocks, $t(6) = 3.04$, $p < .05$; and their number of correct answers was low and constant until the 0 trial block and then increased, as reflected in the increase in percent correct from the –4 to –1 trial blocks to the +1 to +4 blocks, $t(6) = 7.22$, $p < .001$.

The types of “other errors” made just before the discovery closely paralleled the types of other errors made earlier when most errors were duplicates. From the last time when three consecutive answers were duplicates to the first time when three consecutive answers were correct, most errors (62%) were wrong in either orientation or size. This was very close to the 63% of errors that were wrong on those dimensions in Session 1.

The shift from a predominance of duplicate errors to a predominance of “other” errors was not due to the particular problems on the 12 trials before discovery having a greater number of duplicate choices. Because discoveries occurred at widely varying sessions and trials within sessions, and because the 12 trials prior to discovery fairly often included trials from the end of the prior session as well as the beginning of the new one, the number of duplicate choices was similar on the trials on which children met the discovery criteria as on others.

It is worth noting that only through examining the change process at this detailed level was it possible to
identify the shift in the predominant type of error shortly before the discovery. Recall that there was no change over sessions in the percentage of incorrect answers that were duplicates. Even when analysis was limited to the learner group, no change in type of errors over sessions was apparent. The reasons why only the trial-by-trial analysis revealed the change were that different children discovered the correct strategy in different sessions, and that the shift from duplicate to “other” errors was limited to a period of about half a session. Thus, the trial-by-trial analysis allowed insight into the change process that would have been impossible without it.

The trial-by-trial analysis also allowed for examination of the types of problems on which the correct approach was discovered. The main question was whether children were more likely to make discoveries on problems with only one duplicate response possibility rather than two. Examination of the 0 trial block (the first time each child advanced three consecutive correct answers) did not support the view that discoveries would occur disproportionately on these problems. The children’s 0 trial block included 67% problems with two duplicate answers, quite similar to the 55% of total problems with that characteristic. Thus, as in previous microgenetic studies, discoveries occurred on a representative subset of problems.

Rate of change. Examining changes in performance over the seven sessions allowed us to determine the session in which children in the experimental group learned. Children in the experimental group learned quite a bit in Sessions 1 through 4; during these sessions, their percentage of correct answers on the matrix completion task doubled, from 25% to 50%, \( t(29) = 3.87, p < .01 \). Percentage of correct answers increased from Session 1 to 2 (25% versus 36%), \( t(29) = 2.75, p < .01 \), and again from Session 3 to 4 (43% versus 50%), \( t(29) = 2.88, p < .01 \). It did not change significantly from Session 4 to 5, \( p > .2 \), or thereafter, remaining at around 50% from Sessions 4 through 7.

All of the learners first met the criterion for discovery during Session 1 (2 children), Session 2 (3 children), or Session 3 (3 children). As can be seen in Figure 5, the discovery triggered a large increase in correct responding; accuracy increased from 20% correct in the 12 trials immediately before children met the criterion to 78% correct in the 12 trials after they met it. Examination of individual children’s performance indicated that each of the learners performed more accurately on the 12 trials after meeting the criterion than on the 12 trials before meeting it. Performance of the 8 learners varied considerably after they met the criterion, however. Three learners were correct on 100% of the next 12 trials; 2 were correct on 70% to 80% of trials; and 2 were correct on 58% and 33%, respectively. These last 2 children did relatively well on the six trials immediately after they met the criterion—83% and 50% correct—which argues against the interpretation that they might have met the criterion by chance. On the next six trials, however, these children generated fewer than 50% correct answers. Thus, as in previous microgenetic studies, discovery of a new approach did not guarantee consistent reliance on the approach thereafter.

Breadth of change. To test whether children in the experimental group generalized their learning to conservation problems, differences in conservation learning were examined among children whose matrix completion performance led to their being classified as learners, nonlearners, or precocious. As shown in Figure 5, children who were classified as learners on the basis of their matrix completion performance improved from 25% to 62% correct conservation responses from Session 1 to Session 7, \( t(7) = 9.00, p < .001 \). In contrast, children classified as nonlearners on the basis of their matrix completion performance did poorly in both Session 1 and Session 7 on the conservation problems (26% and 30% correct, \( ns \)), and those classified as precocious did well on conservation in both sessions (60% and 70% correct, \( ns \)). The generalization from learning of matrix completion to conservation could not be explained in terms of the learners having higher IQs than the nonlearners. Among the
25 learners and nonlearners (the children who did not know how to solve matrix completion problems at the beginning of Session 1), the correlation between the increase in correct matrix completion answers between Sessions 1 and 7 and the increase in correct conservation responses over the same period was $r = .58$, $p < .01$. Partialing out IQ scores from the correlation reduced it only slightly, $r = .54$, $p < .01$. Thus, acquisition of matrix completion proficiency generalized to conservation above and beyond the effects of IQ.

For the entire sample of 6-year-olds, performance on conservation and matrix completion problems correlated positively both at Session 1, $r(66) = .37$, and Session 7, $r(66) = .35$, $ps < .01$, with the correlations not differing between the two sessions.

Changes with Age

**Matrix completion.** A one-way ANOVA on percentage of correct matrix completion answers for 6-, 7-, and 8-year-olds in Session 1 revealed increases from 20% for 6-year-olds to 48% for 7-year-olds to 78% for 8-year-olds, $F(2, 132) = 38.17, p < .001$. Accuracy increased significantly between 6 and 7 years and also between 7 and 8 years. As in the microgenetic portion of the study, the large majority of errors were duplicates: 57% versus a chance probability of 31%, $t(125) = 9.98, p < .001$. This percentage did not change with age, although the number of all errors fell sharply between 6 and 8 years.

To examine changes with age in percentage of answers that were correct on each dimension, a 3 (age) $\times$ 4 (dimension) ANOVA was performed. The analysis revealed an effect for age, $F(2, 132) = 32.14, p < .001$, but not for dimension or for the Age $\times$ Dimension interaction. Thus, age-related improvements were similar along all four dimensions.

Also conducted was an Age $\times$ Dimension ANOVA on percentage of explanations following incorrect answers on which children cited each dimension in their explanations. The analysis revealed effects for age, $F(2, 113) = 14.46, p < .001$, dimension, $F(3, 339) = 5.07, p < .01$, and the Age $\times$ Dimension interaction, $F(3, 339) = 2.32, p < .05$. Separate ANOVAs for each dimension revealed that the interaction was due to age-related increases in percentage of explanations citing orientation (31%, 46%, and 70% for 6-, 7-, and 8-year-olds, respectively), but no changes in citations of the other three dimensions.

**Conservation.** The number of trials on which children advanced correct answers and explanations on conservation problems increased with age, $F(2, 132) = 28.59, p < .001$. The 8-year-olds and 7-year-olds were correct more often than were the 6-year-olds (74% and 62% versus 24%, respectively).

Relations between Short-Term and Long-Term Change

The parallel data that were collected regarding changes over age and changes over sessions made it possible to compare microgenetic and age-related change. Appropriate comparison of the two types of change required first identifying a measure of the overall amount of change and an age range that involved a similar amount of change on that measure to the amount of change in the microgenetic portion of the study. The reasoning was that if the amounts of change were not comparable, differences in the patterning of change would be impossible to interpret. Changes in percentage of correct answers provided an intuitively reasonable measure of the overall amount of change. The improvement on this measure between ages 6 and 7 years proved to be similar to the improvement between Sessions 1 and 7 for children in the experimental group. In both cases, percentage of correct answers improved from about 20% to 50%.

The next step was to compare the changes between ages 6 and 7 years with the changes between Sessions 1 and 7 on the variables listed in the first column of Table 2. Variable 1 in the table, percentage correct, was the index of overall change on which the cross-sectional and microgenetic samples were matched; by definition, both showed change on that measure. The next eight variables indicated whether number of answers that were correct on each dimension increased and whether number of explanations that cited each dimension increased. As can be seen, presence or absence of change matched on seven of these eight measures. Variable 10 involved children’s predominant type of error when they answered incorrectly. In all seven sessions, most errors were duplicates. Similarly, both 6- and 7-year-olds generated far more duplicate errors than would have been expected by chance. The absolute percentages of duplicate errors also were very similar: 57% among 6- and 7-year-olds; 59% over the seven sessions.

The final two variables (11 and 12) involved stability over time and transfer to other tasks. Both of these are among the criteria that Piaget (1964) proposed as definitional for developmental change. On both variables, the microgenetic changes fit Piaget’s criteria for true development (i.e., development expected to occur with age in the absence of directly relevant instruction). With regard to stability over time, percentage of correct answers remained identical over the 7- to 8-week period between Sessions 6 and 7 (49% correct). Percentages of answers that were correct on each of the four individual dimensions in Sessions 6 and 7 were all within 2% on the two occasions. The correlation of each child’s percentage of correct an-
The basic logic of this study was that by combining microgenetic and cross-sectional components in a single design, a considerable amount could be learned both about the particular acquisition being studied and about the relation between short-term and long-term change. Therefore, this concluding section first focuses on implications of the present findings for understanding of matrix completion and then on implications for how short-term and long-term changes are related.

Acquisition of Matrix Completion Proficiency

Viewing change on the matrix completion task in terms of the path, rate, breadth, variability, and sources of change allowed us to understand better the growth of matrix completion skills and to place this growth in the context of findings from other microgenetic studies.

Path of change. The present study’s findings, both cross-sectional and microgenetic, replicated major parts of Inhelder and Piaget’s (1964) description of the path that leads to understanding of matrix completion: Skill in solving matrix completion problems was found to grow considerably between ages 6 and 8 years. The most common error on such problems was choosing a duplicate of one of the objects in the problem. A third finding that replicated those of Inhelder and Piaget was that choosing response alternatives with the correct orientation posed greater difficulty than choosing ones with the appropriate form, size, or color. On the other hand, the present study’s cross-sectional findings regarding the path of change, like those of Kingma (1983), differed from Inhelder and Piaget’s in not showing a decline in accuracy between ages 6 and 7. In addition, the microgenetic analyses revealed two aspects of the path of change that had not been described by Inhelder and Piaget or by subsequent investigators. Shortly before discovery of the correct approach, the frequency of duplicate errors decreased and the frequency of other errors increased. In addition, discoveries were made on a representative set of problems, rather than being closely linked to problem characteristics, such as number of duplicate choices available.

The present findings regarding the path of change in matrix classification ability were consistent with previous microgenetic findings in other domains. Analyses of growth of understanding of balance scales, number conservation, and mathematical equality have revealed similar initial periods of consistent reliance on a single erroneous approach, then a period of high variability, and then relatively consistent use of the correct approach (Alibali & Goldin-Meadow, 1993; Graham & Perry, 1993; Siegler, 1995;
Siegler & Chen, 1998). Heightened variability around change points also has been observed in human infants, human adults, rats, and pigeons (Machado, 1997; Neuringer, 1992; Stokes, Mechner, & Balsam, 1999; Thelen & Ulrich, 1991). It thus seems to be a common feature of learning.

The present findings also raised the issue of why children so often chose duplicates of the object that was above or alongside the open square in the matrix. One possible reason for the children’s persistence was that on the eight practice trials before the beginning of Session 1, choosing the response that was identical to the one directly above it was always correct. This early history, however, seems inadequate as a full explanation for why so many children would choose duplicates on most of the 88 feedback trials, as almost all nonlearners did, nor for why the percentage of errors that were duplicates would be so consistent over the seven sessions.

Part of the explanation seems likely to be the tendency of young children to rely on perceptual matches at the expense of relational similarity. This tendency was noted by Piaget (1952) and was a major part of his explanation of why children have difficulty on such tasks as conservation, class inclusion, and matrix completion. It also has been noted by Gentner (e.g., Gentner & Toupin, 1986; Ratterman & Gentner, 1998) in connection with analogical reasoning. On analogical reasoning problems, young children often choose object matches, whereas older children choose relational matches. The present study’s matrix completion problems required children to encode same/different relations for the form, size, color, and orientation of the objects in the top row or left column and then to find the response alternative that would create the identical relations for the objects in the bottom row or right column. This relational encoding probably was difficult for many 6-year-olds to generate. If so, inability to generate a superior alternative may have led to the 6-year-olds’ persistent choices of duplicates, despite the duplicates being consistently incorrect after the practice trials.

Rate of change. Those children who met the criterion for discovery answered correctly on 78% of the next 12 problems they encountered. As in previous microgenetic studies, most children sometimes regressed to previous strategies (usually duplication) after having made the discovery. In the context of other findings from microgenetic studies, however, the rate of adoption of the new matrix completion strategy was at the high end.

Comparing rates of change in this and previous microgenetic studies suggested a distinction between the type of task that elicits relatively rapid change and the type of task that elicits slower change. The tasks that have elicited relatively rapid change include conservation (Siegler, 1995), balance scales (Siegler & Chen, 1998), mathematical equivalence (Alibali & Goldin-Meadow, 1993), and now matrix completion. The discoveries that have been examined on these tasks share two qualities that differentiate them from the discoveries on tasks on which adoption of the new approach has been slower. One is that the new approach results in a far higher percentage of correct answers than do previous strategies, and thus results in far more positive feedback. The other characteristic associated with relatively rapid change is that the new discoveries tend to be logically superior to previous ones. For example, relying on the type of transformation to solve a conservation task is logically superior to relying on the length of the rows.

New strategies that have been adopted more gradually have consistently involved less dramatic changes in accuracy. In some cases, such as preschoolers’ discovery of the min strategy (counting on from the larger addend), the new approach results in faster performance, but has minimal effects on accuracy (Siegler & Jenkins, 1989). In other cases, such as experimental design strategies, the new strategy allows surer solutions, but the superior consequences of using the new approach take a long time to become evident (Kuhn et al., 1995; Sch plausible for why the percentage of errors that were duplicates would be so consistent over the seven sessions.

Part of the explanation seems likely to be the tendency of young children to rely on perceptual matches at the expense of relational similarity. This tendency was noted by Piaget (1952) and was a major part of his explanation of why children have difficulty on such tasks as conservation, class inclusion, and matrix completion. It also has been noted by Gentner (e.g., Gentner & Toupin, 1986; Ratterman & Gentner, 1998) in connection with analogical reasoning. On analogical reasoning problems, young children often choose object matches, whereas older children choose relational matches. The present study’s matrix completion problems required children to encode same/different relations for the form, size, color, and orientation of the objects in the top row or left column and then to find the response alternative that would create the identical relations for the objects in the bottom row or right column. This relational encoding probably was difficult for many 6-year-olds to generate. If so, inability to generate a superior alternative may have led to the 6-year-olds’ persistent choices of duplicates, despite the duplicates being consistently incorrect after the practice trials.

Rate of change. Those children who met the criterion for discovery answered correctly on 78% of the next 12 problems they encountered. As in previous microgenetic studies, most children sometimes regressed to previous strategies (usually duplication) after having made the discovery. In the context of other findings from microgenetic studies, however, the rate of adoption of the new matrix completion strategy was at the high end.

Comparing rates of change in this and previous microgenetic studies suggested a distinction between the type of task that elicits relatively rapid change and the type of task that elicits slower change. The tasks that have elicited relatively rapid change include conservation (Siegler, 1995), balance scales (Siegler & Chen, 1998), mathematical equivalence (Alibali & Goldin-Meadow, 1993), and now matrix completion. The discoveries that have been examined on these tasks share two qualities that differentiate them from the discoveries on tasks on which adoption of the new approach has been slower. One is that the new approach results in a far higher percentage of correct answers than do previous strategies, and thus results in far more positive feedback. The other characteristic associated with relatively rapid change is that the new discoveries tend to be logically superior to previous ones. For example, relying on the type of transformation to solve a conservation task is logically superior to relying on the length of the rows.

New strategies that have been adopted more gradually have consistently involved less dramatic changes in accuracy. In some cases, such as preschoolers’ discovery of the min strategy (counting on from the larger addend), the new approach results in faster performance, but has minimal effects on accuracy (Siegler & Jenkins, 1989). In other cases, such as experimental design strategies, the new strategy allows surer solutions, but the superior consequences of using the new approach take a long time to become evident (Kuhn et al., 1995; Schauble, 1996). In yet other cases, such as memorial and attentional strategies, the new approach could somewhat enhance accuracy if it were used effectively, but initial effectiveness tends to be limited by utilization deficiencies (Bjorklund, Coyle, & Gaultney, 1992; Coyle & Bjorklund, 1997; Miller & Aloise-Young, 1996). Thus, it seems highly likely that one determinant of the rate of change is the amount of improvement in accuracy that use of the new strategy brings.

The rate of change also may be influenced by whether the logic underlying the new strategy is superior. At present, this is less certain, however, because the tasks on which the new strategy was logically superior have been a subset of the tasks on which the new strategy generated much more accurate performance. Identifying situations in which logically inferior and superior strategies would yield similarly accurate performance, and observing faster uptake of the logically superior approach, would provide the needed evidence. For example, teaching some children to solve class inclusion problems by counting the number of objects in the superordinate and larger subordinate classes, teaching others to solve the same problems through logical deduction, and then observing the frequency of use of the new strategy in the two groups could address this issue.
Breadth of change. The 6-year-olds’ learning of matrix completion transferred not only to novel matrix completion problems but also to conservation problems. Those children whose matrix completion abilities advanced over the course of the seven sessions showed substantial increases in understanding of conservation. Those children whose matrix completion abilities did not advance did not show increased understanding of conservation. The generalization was not attributable to children with higher IQs making greater progress on both matrix completion and conservation; partialing out IQ left the effect intact. This finding is consistent with Inhelder and Piaget’s (1964) view that ability to consider multiple dimensions simultaneously is essential to understanding of both tasks. It also is consistent with Inhelder, Sinclair, and Bovet’s (1974) finding that training in class inclusion, another task that requires simultaneous consideration of multiple dimensions, transfers to understanding of conservation. Thus, experiences that lead children to consider multiple dimensions of a task, rather than only one, may help them generalize to other tasks that require them to overcome a focus on a single salient dimension.

Variability of change. Microgenetic studies allow examination of variability at several levels. One level is distal variables—qualities that children bring with them to the experimental situation: age, gender, ethnicity, IQ, and so on. In the present study, the distal variable IQ was a moderately strong predictor of learning on the matrix completion task. This was consistent with the findings of Johnson and Mervis (1994), who also found that IQ predicted learning.

A second kind of variability involves individual differences in learning patterns in the experimental situation. Some children showed accurate performance from Session 1 onward; other children started with low levels of accuracy but became accurate sometime within the first three sessions; yet other children never learned to solve the matrix completion problems accurately. Identifying these three groups allowed us to focus analyses of change on children in the learner group; that is, on the children who showed substantial change. The ability to understand the path, rate, and breadth of change was greatly facilitated by being able to identify these children.

A third type of variability involves trial-by-trial changes in performance. The trial-by-trial analyses indicated that children’s predominant error changed roughly a dozen trials before they made the discovery. Before that period, they usually chose duplicates; after it, they chose arbitrarily among the response alternatives. These trial-by-trial analyses also revealed that even after children made the discovery, they sometimes regressed to less sophisticated approaches.

The three types of analysis of variability were mutually informative. Identifying the learner, nonlearner, and precocious patterns made it possible to focus the trial-by-trial analyses on the learners. Similarly, the IQ data provided convergent validity for the distinction among learners, nonlearners, and precocious children; precocious children had higher IQ’s than did nonlearners, with learners in between. Thus, examination of all three types of variability provided a more comprehensive portrait of change than could examination of any one of them alone.

Sources of change. Children in the experimental condition showed greater gains between Sessions 1 and 7 on matrix completion problems than did children in the control condition. The present design did not allow us to separate the contributions of three components of the experimental procedure: experience solving matrix completion problems, feedback concerning the correct answer on them, and efforts to explain why the correct answer was correct. The design, however, did allow us to conclude that together these components were sufficient to produce both learning and transfer. This finding adds matrix completion to the increasing set of tasks on which this procedure has been shown to produce positive effects: number conservation (Siegler, 1995), balance scales (Pine & Messer, 2000), shadow projection (Manz, 2001), biology (Chi et al., 1994), and computer programming (Bielaczyc et al., 1995).

Relations between Short-Term and Long-Term Change

Results of the present study also allow us to compare short-term and long-term change both at the level of behavioral phenomena and at the level of underlying processes.

Behavioral characteristics of short-term and long-term change. To determine how closely behavioral changes produced by a microgenetic manipulation resembled changes that occur with age, parallel data were obtained on 10 measures of performance. On two additional measures, data were obtained from the microgenetic portion of the study and compared with the pattern of change said by Piaget to represent true development. Microgenetic and age-related change proved to be quite similar. On all but one of the measures that could be assessed in both microgenetic and cross-sectional contexts, significant changes were present either in both contexts or in neither. In addition, the two changes that could only be measured in the microgenetic context both met the criteria for developmental change proposed by Piaget (1964). The specificity of the matches was impressive. For exam-
parallels between microgenetic and age-related change perspective, it is especially encouraging that extensive in the everyday environment. From the present perspective, it is especially encouraging that extensive parallels between microgenetic and age-related change are, or are not, similar.

One limitation of the present comparison of short-term and long-term change involved the use of cross-sectional rather than longitudinal data. Longitudinal data on age-related changes would have allowed for direct comparisons of the stability of changes over short- and long-time scales, whereas the present cross-sectional data did not. Such longitudinal data also would have allowed for the ability to control statistically for variations in the initial knowledge of each child, which may have strengthened some of the findings regarding changes with age. Thus, future studies are needed that employ the same basic paradigm but examine age-related changes longitudinally rather than cross-sectionally.

An encouraging implication of the present findings for microgenetic research is that if a change emerges in a microgenetic study, it is likely to be paralleled by an age-related change. On every measure that revealed microgenetic change, there was a corresponding change with age. It thus seems reasonable to hypothesize that phenomena that have emerged consistently in microgenetic studies, but cannot easily be examined in studies of age-related change, also characterize age-related change in the everyday environment. These phenomena include variability immediately prior to discoveries, short-lived transition strategies, and less than instantaneous generalization of new strategies. Longitudinal studies with frequent sampling of strategy use but without any instructional manipulation could be used to test this hypothesis.

It is important to note that the generality of the parallels found in the present study between microgenetic and age-related change is unknown. Our operating assumption was that when the overall amount of microgenetic and age-related change is similar, the detailed patterning of change also is likely to be similar. The present findings were consistent with this assumption. The range of conditions over which the assumption holds true is an empirical question, however. It seems likely that the more closely the microgenetic manipulation corresponds to a denser presentation of the types of experiences that are the source of change in the everyday environment, the more similar that age-related and microgenetic change will be. Indeed, Kuhn (1995) defined microgenetic manipulations as involving exercise of existing strategies, a mechanism that definitely produces cognitive growth in the everyday environment. From the present perspective, it is especially encouraging that extensive parallels between microgenetic and age-related change also arise from procedures like those used in the present study, in which children are asked questions intended to encourage construction of new reasoning strategies. More generally, applying the present type of design with varying types of microgenetic manipulations should allow researchers to establish the conditions under which microgenetic and age-related changes are, or are not, similar.

Processes underlying change. Short-term and long-term behavioral changes clearly have a great deal in common. The same seems likely to be true of the processes that underlie the behavioral changes. To be specific, the processes that produce changes over days or weeks also are likely to be involved in producing changes over years, although the long-term changes may reflect other processes as well.

The detailed descriptions of change yielded by microgenetic studies can help to elucidate the processes that generated them. For example, the present microgenetic data indicated that rejecting existing approaches and generating superior new approaches are separate processes. Children shifted away from choosing a duplicate of one of the objects in the matrix roughly a dozen trials before they discovered the correct strategy. During those dozen trials, they chose in a seemingly haphazard way among the six response alternatives. Then they discovered the new, correct approach and relied on it relatively consistently thereafter. Similar temporal separations between rejection of approaches that yield consistently wrong choices and discovery of superior alternatives have emerged in previous microgenetic studies of learning about balance scales (Siegler & Chen, 1998), number conservation (Siegler, 1995), and mathematical equivalence (Alibali, 1999). These findings suggest that in some situations, giving up on flawed approaches activates strategy discovery mechanisms that generate superior new strategies fairly quickly.

A final question raised by the present study is “What cognitive processes are unique to long-term change, and in what way does behavior produced by them differ from behavior produced by short-term change processes?” The present study’s findings do not answer this question. By documenting the extensive similarities between short-term and long-term cognitive change, however, the present findings raise the question in clear relief.

ACKNOWLEDGMENTS

Much of the research reported in this article was made possible by funding from the Spencer Foundation and National Institutes of Health Grant 19011.
REFERENCES


Machado, A. (1997). Increasing the variability of response sequences in pigeons by adjusting the frequency of


