

Overcoming the Positive-Capture Strategy in Young Children: Learning About Indeterminacy

David Klahr and Zhe Chen

Two experiments were conducted to examine whether and how 4- and 5-year-olds learn to distinguish determinate from indeterminate evidence. Children were asked to decide whether various patterns of evidence were sufficient to reach unambiguous conclusions. This study replicated the finding that young children tend to use a strategy that, although generally successful, fails on evidence patterns in which a single positive instance co-occurs with an unexplored source of evidence. Experiment 1 demonstrated that this positive-capture strategy is deeply entrenched, even in a meaningful, pragmatic context. With a microgenetic design, Experiment 2 revealed that young children are capable of replacing the positive-capture strategy with a correct strategy when they are exposed to various analogous tasks in several training sessions.

The ability to distinguish between determinacy and indeterminacy is a fundamental aspect of higher order cognition in many domains, including causal reasoning, decision making, and scientific discovery. In each domain, it is necessary to distinguish between situations in which evidence is sufficient for drawing conclusions and situations in which it is not. Given the centrality of this distinction, it is important to understand the origin of the ability, its developmental course, and factors that facilitate its acquisition.

Inhelder and Piaget (1958) first investigated the question in terms of the development of possibility and necessity, and since then the topic has been investigated under several other rubrics, including possibility versus impossibility, certainty versus uncertainty, sufficiency versus insufficiency, and

knowing versus guessing (Braine & Romain, 1983; Byrnes & Beilin, 1991; Fabricius, Sophian, & Wellman, 1987; Fay & Klahr, 1996; Johnson & Maratsos, 1977; Johnson & Wellman, 1980; Miscione, Marvin, O'Brien, & Greenberg, 1978; Moore, Bryant & Furrow, 1989; Perner, 1991; Ruffman & Olson, 1989; Somerville, Hadkinson, & Greenberg, 1979; Wimmer, Hogrefe, & Perner, 1988). These studies had different emphases and used a variety of methods for investigating children's understanding of indeterminacy. Consequently, a consistent picture of its developmental course has yet to emerge. In contrast to Piaget's conclusion that such understanding is not attained until the stage of formal operations (i.e., around 11 years), several studies have shown that children as young as 6 years old can distinguish between determinate and indeterminate problems (Byrnes & Overton, 1986; Falmagne, Mawby, & Pea, 1989; Horobin & Acredolo, 1989; Sodian, Zaitchik, & Carey, 1991; Wollman, Eylon, & Lawson, 1979). For example, in Somerville et al.'s (1979) study, children were told that "the boy lives in the house with the blue table in the front yard." When there were two such houses, 6-year-olds, but not 5-year-olds, decided that they needed further information to specify the correct location. However, in tasks requiring the generation of such discriminating experiments, 9-year-olds perform poorly (Chen & Klahr, 1999; Klahr, 2000), and not until age 12 do children consistently distinguish between determinate and indeterminate situations in complex contexts (Acredolo & Horobin, 1987; Byrnes & Overton, 1986; Morris & Sloutsky, 2002; Osherson, 1975; Piérait-Le Bonniec, 1980; Scholnick & Wing, 1988; Wollman et al., 1979).

David Klahr, Department of Psychology, Carnegie Mellon University; Zhe Chen, Department of Human and Community Development, University of California, Davis.

This work was supported in part by grants from the National Institute of Child Health and Human Development (NICHD; HD25211) to the first author and from the NICHD (HD19011) to the second author. Thanks to the children, parents, principal, and teachers at the Carnegie Mellon University Children's School, the Temple Children's Center, the Children's Center of Pittsburgh, and the Shady Lane School for their participation and cooperation. We also thank Jennifer Schnakenberg for her insightful suggestions at several stages of the project, Anne Siegel and Jolene Watson for data collection and analysis, Amy Masnick and Lara Triona for comments on an earlier draft, and Audrey Russo for assistance with the final touches. Finally, we thank the action editor, Douglas Frye, and an anonymous reviewer for their insightful and constructive comments on an earlier version of this paper.

Correspondence concerning this article should be sent to David Klahr, Department of Psychology, Carnegie Mellon University, 500 Forbes Avenue, Pittsburgh, PA 15213. Electronic mail may be sent to Klahr@cmu.edu.

© 2003 by the Society for Research in Child Development, Inc.
All rights reserved. 0009-3920/2003/7405-0003

One reason for this lack of resolution comes from the relatively large grain size at which the process of evidence interpretation has been studied. That is, most investigators have taken a binary view of indeterminate versus determinate situations, rather than a nuanced approach that is sensitive to particular kinds of indeterminate evidence patterns. Fay and Klahr's (1996) analysis of preschoolers' responses to different types of evidence patterns revealed substantial differences in their ability to interpret correctly different patterns of indeterminate evidence. Because the two experiments described in this article are based on the Fay and Klahr approach, we summarize their procedure, use their notation to characterize the issues to be addressed, and summarize their unexpected findings as a prelude to the present study.

Fay and Klahr (1996) studied preschoolers' understanding of indeterminacy by presenting them with a target object (e.g., a necklace made from red beads) and a set of boxes, each of which contained only one type of bead. At the outset of each trial, all boxes were closed. Then they were opened sequentially. Before the opening of each box, children were asked whether and why they "knew for sure" or "would have to guess" about which box was used to construct the necklace. Consider, for example, Problem A in Table 1. Here, the target is a necklace made entirely of red beads, and there are three boxes, containing, respectively, blue, red, and red beads. Our terminology is as follows: (a) *Problem* refers to a set of boxes (whether open or closed) and their contents. (b) *Evidence pattern* (or sometimes just *pattern*) refers to the specific distribution of matching, not matching, and closed boxes. Thus, for each three-box problem, four distinct evidence patterns

are generated from the initial state when all boxes are closed to the final state when they are all open. (c) *Evidence type* designates a set of patterns sharing the same logical properties (e.g., visibly indeterminate for $+-+$ and $-++$). (d) The symbols $?$, $+$, and $-$ denote, respectively, a closed box, an open box that matches the target, and an open box that does not match. For the problem described earlier, the sequence of evidence patterns—depicted in successive rows in the top half of Table 1—is as follows: before any boxes are opened: $???$; after the first box, containing nonmatching beads, is opened: $-??$; after the second box, with red beads, is opened: $-+?$; and after all boxes are opened: $-++$. Note that this problem is indeterminate throughout. That is, there is never sufficient evidence to eliminate the uncertainty about which box was used to construct the necklace of red beads. In contrast, Problem B in Table 1 also starts out with indeterminate evidence patterns, but it becomes fully determinate once the final box is opened.

Fay and Klahr (1996) used a mix of different problems—some that ended with a determinate outcome and some that did not. Their series of three experiments clearly indicated that preschoolers' understanding of indeterminacy was highly dependent on the type of evidence. For most evidence types, children did well. More specifically, for visibly determinate evidence (e.g., a pattern such as $--+-$) children's performance was nearly perfect, and for visibly indeterminate evidence (e.g., patterns such as $-+-+$ or $++++$) children responded correctly on approximately 70% of the trials.

However, there was one evidence type—labeled as positive and hidden in Table 1—for which only

Table 1
Patterns of Evidence for an Indeterminate Problem (A) and a Determinate Problem (B), With a Red Target

Problem	Box contents and exposure level			Evidence pattern			Correct response	Evidence type
	Box 1	Box 2	Box 3					
A	Blue	Red	Red	$?$	$?$	$?$	Guess	Hidden
	Blue	Red	Red	$-$	$?$	$?$	Guess	Negative and hidden
	Blue	Red	Red	$-$	$+$	$?$	Guess	Positive and hidden
	Blue	Red	Red	$-$	$+$	$+$	Guess	Positive and visible (indeterminate)
B	Red	Green	Yellow	$?$	$?$	$?$	Guess	Hidden
	Red	Green	Yellow	$+$	$?$	$?$	Guess	Positive and hidden
	Red	Green	Yellow	$+$	$-$	$?$	Guess	Positive and hidden
	Red	Green	Yellow	$+$	$-$	$-$	Know	Positive and visible (determinate)

Note. For each problem, the sequential opening of boxes from left to right is indicated by successive rows of the table. Problems start with all three boxes closed and end with all three boxes open. (Boldface cells correspond to closed boxes.) Fay and Klahr (1996) used four-box problems in Experiments 2 and 3. In this example we use only three boxes, corresponding to the materials used in the present study.

about 20% of children's responses were correct. Such patterns include at least one closed box and a single positive instance (e.g., $+-?$). On such problems, children incorrectly responded "know" when they should have said "guess" approximately 80% of the time.

In an analysis of the strategies used by individual children, Fay and Klahr (1996, Experiment 2) found that more than 60% of kindergarten children followed a strategy in which they responded correctly to all but the positive and hidden type of pattern. Fay and Klahr called this the positive-capture strategy because the single positive instance seemed to capture children's attention and, in effect, blind them to the fact that the unexplored option might yet render the problem indeterminate.

The ability to distinguish determinate from indeterminate evidence is a prerequisite to appreciating the logic of evidence-theory relations and confounded versus unconfounded experimental designs. The positive-capture strategy shares the same roots of what Kuhn and colleagues have called inclusion errors in scientific reasoning (e.g., Kuhn, Garcia-Mila, Zohar, & Andersen, 1995). The inclusion errors involve the attribution of causal relation of a variable that covaries with the outcome on only a single occasion. Younger children are particularly likely to generate this kind of error, but 11- and 14-year-olds, and even adults, often use the incorrect single-instance inclusion inferences (Kuhn et al., 1995). Thus, a form of the positive-capture strategy appears to cause difficulty in scientific reasoning well into adulthood.

Given the consequences of failing to recognize indeterminate situations in scientific reasoning, it is important to investigate whether and how young children eventually replace the positive-capture strategy with a more advanced strategy. There are two commonly used methods of exploring the robustness of children's failures to exhibit adult levels of performance on various tasks. One approach is to recast formal, abstract problems in everyday, pragmatic contexts by introducing meaningful content. Such studies (e.g., Harris & Nunez, 1996; Light, Blaye, Gilly, & Giroto, 1989) have revealed the powerful role of pragmatic context in children's inductive reasoning.

The other common approach to improving performance is to provide children with extensive experience or direct instruction in the domain, and then to measure the rate and extent of knowledge acquisition and retention. The long history of conservation training studies exemplifies this approach (Brainerd & Allen, 1971; Brainerd & Brai-

nerd, 1972; Field, 1987; Gelman, 1969). More recent microgenetic studies not only examined whether children can acquire a rule, solution, or strategy but also addressed the issue of how change occurs with age and experience. The microgenetic approach involves observing children's performance and strategy change on a trial-by-trial basis as children encounter experiences that might promote learning within particular domains (Kuhn, 1995; Miller & Coyle, 1999; Siegler, 1995, 1996; Siegler & Crowley, 1991).

The primary goal of the present research was to explore the conditions under which the positive-capture strategy comes to be replaced with a more advanced strategy for evaluating evidence. We used both of the approaches described earlier. In Experiment 1 we examined whether and how young children's performance could be improved by presenting tasks in a richer and more socially meaningful context, and in Experiment 2 we provided extensive problem-solving experience and explicit feedback to facilitate learning.

In Experiment 1 we extended Fay and Klahr's (1996) study to pragmatic contexts in which the result of resolving the indeterminacy would have meaningful social consequences: the assignment of credit or blame. Children were asked to evaluate evidence either in a neutral condition similar to those in Fay and Klahr or in a contextualized condition in which discovery of a determinate cause would result either in a positive outcome (the attribution of social credit) or a negative outcome (the attribution of social blame). The goal of this experiment was to determine whether the social outcome resulting from children's judgment of determinacy or indeterminacy would motivate them to be more mindful and careful in evaluating evidence patterns and thus reduce positive-capture errors. The first experiment also examined whether children would learn from experience and indirect feedback.

The second experiment was a microgenetic study designed to examine whether, with extensive experience on multiple trials and several analogous tasks, accompanied by direct, explicit feedback, young children would abandon the positive-capture strategy and improve their understanding of determinacy and indeterminacy.

Experiment 1

The primary aim of Experiment 1 was to determine whether preschoolers' use of the positive-capture strategy could be reduced by embedding

indeterminacy problems in pragmatic contexts. Evidence patterns were presented in situations in which indeterminacy judgments would have meaningful social consequences: the assignment of credit or blame to another child for either good performance or failure to follow rules. We used a credit-blame cover story for three reasons. First, as noted earlier, many investigators have found that when problems are couched in meaningful contexts, pragmatic reasoning schemas can lead to improved performance (cf. Cheng & Holyoak, 1985). Second, the research on scientific reasoning has demonstrated that, for both adults and children, the decision to search for additional evidence before accepting a hypothesis is strongly influenced by the desirability of the hypothesis under consideration (Klahr, 2000; Klahr, Fay, & Dunbar, 1993; Kunda, 1990; Tschirgi, 1980). Third, the literature on social development suggests that preschool children are highly motivated to engage in discussions of and ratings for situations involving credit and blame (Dunn, 1988; Nelson-le Gall, 1984; Schleifer, Schultz, & Lefebvre-Pinard, 1983).

In this experiment children were asked to decide whether various evidence patterns constituted determinate or indeterminate situations. In the no-context condition, the evidence patterns were abstract (as in Fay and Klahr, 1996), whereas in the context condition, a cover story explained that the outcome of the evidence-evaluation process would result in either credit or blame being ascribed to another—fictitious—child. More specifically, in the context condition, discovery of a determinate pattern would result in the fictitious child's receiving credit (praise for producing "an exceptionally good drawing") in some cases and blame (having to "clean up a mess") in others. We conjectured that children would prefer indeterminate outcomes in blame situations (because they could avoid being in the undesirable position of attributing blame to another child) and determinate outcomes in credit situations (because they would be in the desirable position of attributing credit). These preferences, in turn, should result in a more complete analysis of evidence in blame situations than in credit (or neutral) situations. This would produce fewer positive-capture errors when the outcome is blame (i.e., children would correctly say "don't know") and more positive-capture errors when the outcome is credit (i.e., children would incorrectly say "know"). On the other hand, research suggesting that children tend to attribute more blame for bad deeds than credit for good deeds (Fincham, 1985) predicts the opposite effect of credit and blame on indeterminate pro-

blems. That is, in many contexts children appear to view punishment for misdeeds as more important than rewards for good deeds (Hamilton, Blumenfeld, & Kushler, 1988). More generally, none of the many investigations of how children make such attributions has examined how evidence-evaluation processes interact with the assignment of credit and blame. Thus, Experiment 1 represents an initial exploration into this area, as well as an investigation of the robustness of the positive-capture phenomenon.

Method

Participants

Participants were 98 children between 4.0 and 6.3 years of age. They were recruited via requests for parental permission from four local preschools. In each school, children were randomly assigned to one of two conditions (neutral vs. context). For purposes of later analysis, a median split on age was used to classify children as 5-year olds (M age = 5.4 years, $range = 4.9$ – 6.3 years) and 4-year-olds (M age = 4.5 years, $range = 4.0$ – 4.8 years). Approximately the same proportions of girls and boys were in each group (43% and 57%, respectively).

Materials

Problems were three-box variants of the four-box problems used in Fay and Klahr (1996). Thirty-six lidded boxes, each containing a magic marker, were used, as well as four small tables and monochrome pictures drawn with markers. Materials were used in the different phases as follows.

Familiarization tasks. Six boxes, six markers (three blue, two red, one green) and a piece of white paper were used.

Practice tasks. Tasks consisted of six boxes mounted in sets of two on a white foam board and six markers (yellow and green, purple and purple, and brown and brown). There were one determinate and two indeterminate problems.

Test problems. On Day 1, 12 boxes—3 on each of four tables—were used. Each table had cardboard table top covered with wood-grain contact paper, on which was a drawing, scribble, or design. There were pictures of three boys or three girls at each table, one in front of each box. In Day 2, 12 boxes were used, as in Day 1, but different boxes, markers, drawings or scribbles or designs, and children's pictures were used. We used four orders of eight problems (four determinate and four indeterminate). Children in the context condition received four

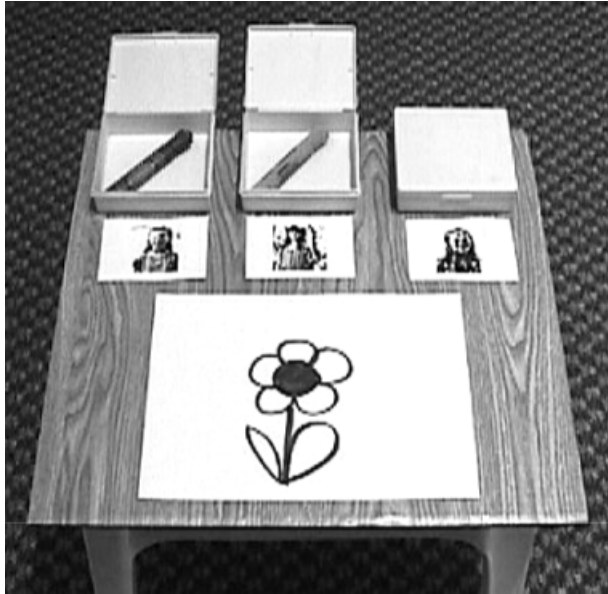


Figure 1. Example of a credit problem with two of the three boxes open (Experiment 1). The goal is to determine which of the three (fictitious) children pictured below each box drew the “pretty” picture so that they can get credit for their work. The first open box contains a marker matching the picture color, the second open box contains a nonmatching marker, and the third box has not yet been opened (corresponding to the $+ - ?$ pattern of evidence shown in Table 3).

credit and four blame problems. A typical problem presentation is shown in Figure 1.

Design

We varied the context by presenting the problems either in a neutral situation—similar to that in Fay and Klahr (1996)—or in a pragmatic context, in which a cover story embedded the problems in social situations in which the resolution of indeterminacy would result in the assignment of either credit or blame. In summary, this was a 2 (condition: context vs. neutral) \times 4 (evidence type) \times 2 (age) mixed design, with evidence type as a within-subjects variable. The credit versus blame variable was nested within the context problems. Finally, problem order was used as a factor in some analyses to explore possible improvement in performance over repeated exposure to indeterminacy problems.

Procedure

Children were tested individually in a room in their school. Three phases were presented: familiarization, practice, and test problems. Children sat at a large table with the experimenter for the familiariza-

tion and practice problems and then walked around the room to the smaller tables for the test problems. In both the practice phase and the test phase, the boxes were first presented closed and then opened one at a time by the experimenter, from the child's left to right.

Familiarization phase. During familiarization, the experimenter showed the child six marker boxes and opened each one. Three boxes contained a blue marker, two contained a red marker, and one contained a green marker, so that children could see that each box always contained only one marker and that it was possible for the same color marker to be in more than one box. Children were invited to draw a small picture or write their name on a piece of white paper, using the markers from the familiarization boxes.

Practice phase. Children were presented with three 2-box problems: one determinate and two indeterminate. The second indeterminate problem was presented only if children had difficulty recognizing that the evidence when both boxes were open in the indeterminate problem meant that they “had to guess.” For each practice problem, children were shown two closed marker boxes mounted on white foam board and a design (an abstract drawing, such as a row of squares) drawn with a single color of marker. Children were told that each box contained one marker and that someone used the marker from just one of the boxes to make the design. They were then asked the test question: “With both boxes closed, do you know for sure which box was used to make this design, or do you have to guess?” (Throughout the practice and experimental phases, the order of “know for sure” and “have to guess” was randomized in the test question.)

Feedback was provided during the practice problems. After children indicated whether they knew for sure or had to guess with both boxes closed, they were told that in such a situation, “You have to guess because both boxes are closed and you can't see inside them.” They were also told that in the game they were playing with the experimenter, there would be times when they would have to guess and other times when they could know for sure. It was emphasized that whenever they did not know for sure they should say that they had to guess.

The experimenter then opened the boxes one at a time, revealing patterns of evidence (e.g., $+ ?$ followed by $++$, or $+ ?$ followed by $+ -$). After each box was opened, the child was asked, “Do you know for sure which box was used to make the design, or do you have to guess?” Children were

also asked how they knew for sure or why they had to guess. At the end of each practice problem, after both boxes were opened and children had responded to them, feedback was given, explaining how the children could know for sure or why they had to guess, for determinate and indeterminate problems, respectively. During the feedback, the experimenter pointed out the color of the design, the color of the marker in each box, and the match between them. By the end of the practice task, children's understanding of the basic aspects of the procedure was well established. Most important, they demonstrated that they were comfortable with giving both know and guess responses.

Experimental phase. The procedure for the experimental phase is summarized in Table 2. This phase consisted of the four 3-box problems shown in Table 3 presented in several different orders. Children were told they were going to play a pretending game with the marker boxes. The experimenter and child then walked to the first of four small tables placed throughout the room. On each table were three marker boxes, and in front of each marker box was a black and white photograph of a child of the same gender as the subject (see Figure 1). In the contextualized condition, on the table there was also either a piece of paper with a picture on it (credit problems) or a scribble made directly on the table

top (blame problems), whereas in the neutral condition, there was a paper with an abstract design on it, similar to those used in the practice problems. Children in the context condition received problems in two credit and two blame contexts in one of five different counterbalanced orders.

At each table, the experimenter pointed out the boxes and photographs to the child and explained about the drawing, mess, or design and indicated what the child should do, as detailed in Table 2. Before any boxes were opened, and after each box was opened, children were asked the test question, "Do you know for sure who made this picture/mess/design, or do you have to guess?" They were then asked to explain their response and were asked one of two probe questions—"How do you know for sure?" or "Why do you have to guess?"—depending on their answer to the test question. The experimenter then opened the boxes sequentially, asking the test and probe questions after each box was opened. No feedback was provided.

During their first day in the experiment, children completed the familiarization and practice phases and the first four problems of the experimental phase. On the next day, children were reminded of the pretend classroom context, using the same description as earlier and then completed the remaining four problems of the experimental phase.

Table 2
Procedure for Experimental Phase of Experiment 1

Context condition		
Credit	Blame	Neutral condition
<ul style="list-style-type: none"> When we get to this table, you see that one of the boys/girls left his/her picture on the table. It looks like he/she worked really hard on his/her picture. I think it's a great picture, and I want to hang it up in the classroom so that the other children can see it. I want to find out which boy/girl made this picture so that I can put his/her name on it and hang it up. Can you help me figure out which boy/girl made this picture? 	<ul style="list-style-type: none"> When we get to this table, you see that one of the boys/girls scribbled on the table top with his/her marker. It looks like he/she didn't follow the rules. I'm really annoyed about the mess, and I want to get it cleaned up. I want to find out which boy/girl made this mess so that I can tell him/her to clean it up. Can you help me figure out which boy/girl made this mess? 	<ul style="list-style-type: none"> When we get to this table, you see that one of the boys/girls made a design with his/her marker. I wonder who made it. Can you help me figure out which boy/girl made this design?

Note. After a brief familiarization and practice phase, children were told the following: This is going to be a pretending game, and this is our pretend classroom. I'm the teacher in this classroom, and you're going to be my helper. Right now the other children from this classroom are out on the playground, and we're going to go around and see if everything in the classroom is OK. In this pretend classroom, the children were drawing with markers at their tables earlier in the day. There were three children at each table, and each child had their very own marker box, and inside each box was one marker, like in the boxes we were just playing with. Each child made their own drawing, and they only used the marker from their own box. Now we're going to go around to the tables in this classroom and see what we find at each table.

Table 3
Types of Problems, Patterns of Evidence, and Their Frequencies of Occurrence (Experiment 1)

Panel A. Problem types and evidence patterns					
Problem type ^a					
	Indeterminate	Indeterminate	Determinate	Determinate	
Evidence patterns as boxes are opened and contents revealed	???	???	???	???	
	+??	-??	+??	-??	
	+ - ?	- + ?	+ - ?	- - ?	
	+ - +	- + +	+ - -	- - +	
Panel B. Frequency of different types of evidence patterns					
Evidence pattern	Correct response	Pattern frequency	Evidence type	Type frequency	Number of exposures to each type over Days 1 and 2
???	Guess	4	Hidden	4	8
-??	Guess	2	Negative and hidden	3	6
- - ?	Know ^b	1			
- ? ?	Guess	2	Positive and hidden	5	10
+ - ?	Guess	2			
- + ?	Guess	1			
+ - +	Guess	1	Positive and visible	4	8
- + +	Guess	1			
- - +	Know	1			
+ - -	Know	1			

^aThe four problem types in Panel A were presented in several different orders in each condition.

^bThis is determinate, by inference, because at least one match is guaranteed by the rules of the game.

Results

Two major independent variables were measured. The first measure involved children's response to the know-guess question ("Do you know for sure who made this picture/mess/design, or do you have to guess?") on each trial. Responses were coded as correct or incorrect and assigned a score of 1 or 0, respectively. For indeterminate patterns, the correct response was "have to guess," whereas for determinate patterns, the correct answer was "know for sure."

The second measure was the number of children who used an expert rule. The expert designation here is arbitrary and is intended simply as a suggestive label for the children who correctly solved a substantial proportion of different types of problems. As shown in Panel A in Table 3, children saw 16 evidence patterns as the four different problems types went from fully closed to fully open, and each problem was presented twice, yielding a total of 32 responses for each child. Ten of these were responses to positive and hidden evidence patterns. We classified children as expert rule users if they

responded correctly on at least 26 of the 32 total evidence patterns, including at least 8 of the 10 positive and hidden evidence patterns. Although these criteria are arbitrary, the use of an even more restricted criterion yielded similar performance patterns.

Our analysis addresses the following five questions: (a) Are children more likely respond correctly in the context condition than in the neutral condition? (b) Do children in the context condition perform differently when they encounter credit and blame patterns? (c) Do 5-year-olds generate more correct responses than 4-year-olds? (d) Do children experience particular difficulty in solving positive and hidden patterns? (e) Do children improve their performance on the most difficult pattern type—positive and hidden—over trials? The answers to these five questions are organized into three sections: (a) children's mean performance on different types of evidence patterns, (b) the proportion of expert rule users in different conditions at different age levels, and (c) percentage of children responding correctly to the 10 presentations of positive and hidden evidence patterns.

Percentage of Correct Responses in Solving Different Types of Patterns

Figure 2 shows the percentage of correct responses for each evidence type in the neutral and context conditions and at the two age levels. A 2 (condition: neutral vs. context) \times 2 (age: 4- vs. 5-year-olds) \times 4 (evidence type: hidden, negative and hidden, positive and visible, and positive and hidden patterns) analysis of variance (ANOVA) with pattern type as a repeated measure was performed. The analyses yielded a main effect for age, $F(1, 94) = 18.86, p < .0001$, and a main effect for evidence type, $F(3, 282) = 81.25, p < .0001$. Neither a main effect for condition nor any interactions were obtained. The main effect for age reflected the higher performance in 5-year-olds (76% correct) than 4-year-olds (59% correct). Further analyses revealed that children had difficulty in solving the positive and hidden evidence types: Their performance on this type of evidence (39% correct) was significantly lower than on the hidden (77%), negative and hidden (74%), and positive and visible evidence types (81%), $ps < .0001$. Children's performance on positive and visible evidence types was significantly higher than on negative and hidden evidence types, $p < .001$.

The lack of an overall context effect could result from averaging over effects of credit and blame that go in opposite directions. Thus, further analyses were conducted to examine the possible effects for type of context (credit or blame). Given that credit and blame problems were presented only in the context condition, only children in this condition

were included in the analyses. Figure 3 shows 4- and 5-year-olds' performance on credit and blame problems. A 2 (age) \times 2 (context type: credit vs. blame) \times 4 (evidence type: hidden, negative and hidden, positive and visible, and positive and hidden) ANOVA with context type and pattern type as repeated measures was performed. The analyses yielded a main effect for age, $F(1, 45) = 5.51, p < .05$, and a main effect for pattern, $F(3, 135) = 34.79, p < .0001$. Neither a main effect for context type nor any interactions were obtained. Children performed similarly in solving credit and blame problems (67% and 68% correct, respectively).

To explore further the apparent lack of a condition effect, we examined children's performance on the

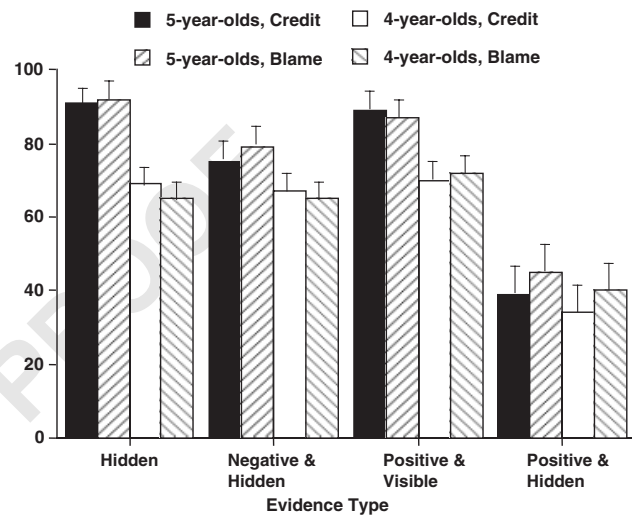


Figure 3. Percentage of correct responses to each type of evidence by context type for context problems, with standard error bars (Experiment 1).

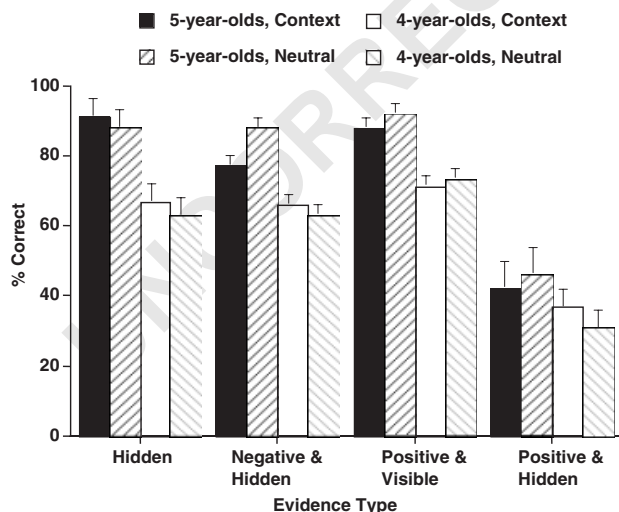


Figure 2. Percentage of correct responses to each type of evidence by conditions and age, with standard error bars (Experiment 1).

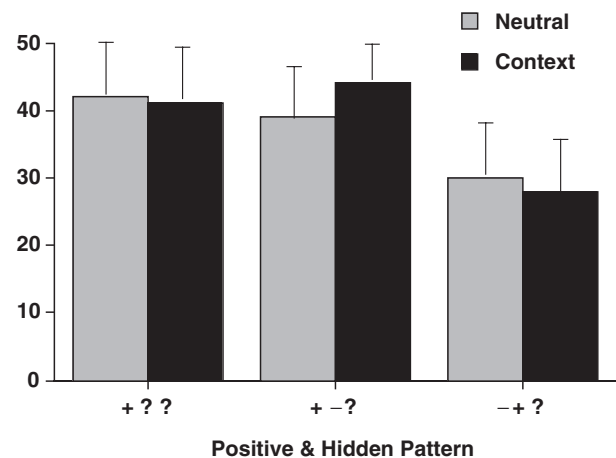


Figure 4. Percentage of correct responses to three positive and hidden evidence patterns, with standard error bars (Experiment 1).

three positive and hidden evidence patterns (+ ? ?, + - ?, and - + ?; see Figure 4.) A 2 (condition: neutral vs. context) \times 2 (age) \times 3 (evidence pattern) ANOVA with evidence pattern as a repeated measure yielded a main effect for evidence pattern, $F(2, 192) = 11.05$, $p < .0001$, but not for condition or age, and it yielded no main effect for any interactions. Children in both conditions and at both ages performed similarly; they had difficulty in solving each of the positive and hidden patterns (41%, 42%, and 29%, for the + ? ?, + - ?, and - + ? patterns, respectively), with substantially poorer performance on the - + ? evidence pattern. Thus, although as noted earlier, 5-year-olds outperformed their younger classmates on the easier evidence evaluation tasks, they showed no advantage when presented with the difficult positive and hidden patterns.

Expert Rule Use

To obtain a more complete picture of children's overall performance we examined individual differences in expert levels of performance. Overall, only 18 of 98 children (18%) met the criteria for being expert rule users. In the neutral condition, 12% and 20% of the 4- and 5-year-olds, respectively, used the expert rule, and in the context condition, 17% and 25% of the 4- and 5-year-olds, respectively, used the rule. In both conditions children rarely exhibited expert rule use, mainly because they performed so poorly on the positive and hidden patterns (recall that the expert rule criterion includes correctly responding to at least 8 of the 10 positive and hidden patterns).

Performance on Positive and Hidden Evidence Types Over Trials

Although Experiment 1 was not designed as a learning study, it is possible that children did learn over the course of repeated exposure to different types of problems. Of particular interest is whether they improved their performance in evaluating the difficult positive and hidden patterns. As noted earlier, children encountered 10 such patterns distributed over 32 trials (see Table 3). However, a simple temporal analysis could be misleading because performance might be influenced by the differential mix of harder and easier variants of positive and hidden patterns early and late in the overall sequence of trials. For example, if children received the - + ? pattern in an earlier phase more frequently than in a later phase, the improvement would not necessarily reflect children's learning with experience because the - + ? pattern is more

difficult to solve than the other positive and hidden patterns. Although the frequencies of this pattern differed in order positions, it is comparable between the 1st, 3rd, and 10th trials: 43%, 47%, and 47% of the older children encountered the - + ? pattern as the 1st, 3rd, and 10th patterns, respectively, and 51%, 39%, and 39% of the younger children received this evidence pattern as the 1st, 3rd, and 10th patterns, respectively. An analysis based on these three occurrences shows that older children increased their performance from 20% correct (1st occurrence) to 39% (3rd occurrence) to 51% (10th occurrence), whereas younger children did not improve their performance (31%, 35%, and 37% on the 1st, 3rd, and 10th occurrences, respectively). A 2 (age) \times 3 (1st, 3rd, vs. 10th trial) ANOVA with trial as a repeated measure yielded a main effect for trial, $F(2, 192) = 5.45$, $p < .005$, no main effects for age, but a marginally significant interaction between trial and age, $F(2, 192) = 2.41$, $p = .09$. Thus, there is preliminary evidence for learning with experience in 5-year-olds but not in 4-year-olds.

Discussion

Children's responses to the different evidence patterns replicated the Fay and Klahr (1996) study and yielded a robust positive-capture effect. Both 4- and 5-year-olds tended to respond "know" whenever they encountered a positive and hidden pattern. Contrary to our expectations, positive-capture errors were equally likely in the neutral and context conditions, and within the context condition, equally likely for credit and blame problems. Although the present manipulation represents a step toward a more pragmatic context, it is possible that our operationalization of meaningfulness was not sufficiently robust to make children treat the contextualized cover story any differently than they treated the no-context problems. That is, discovering something in the evidence pattern that would "help the teacher to find a (hypothetical) child to praise or blame" may not have meant much to our participants. Another possibility is that positive capture is a fairly robust, albeit inadequate, strategy, unaffected by reasoning about evidence in a richer social context. Further studies would be necessary to resolve this question. However, Experiment 1 did suggest that, of the two approaches mentioned earlier (i.e., contextualization and training) for enhancing children's performance on recalcitrant error-prone strategies, the training approach might be applicable to the positive-capture strategy. We turn to that issue next.

Recall that because Experiment 1 was not designed as a training study it contained no explicit

feedback about the correctness of judgments of determinacy or indeterminacy. Nevertheless, the sequential opening procedure for each set of boxes did provide implicit feedback in some cases. Previous studies in several domains (e.g., Bjorklund & Green, 1992; Klahr, 2000; Siegler, 1995; Siegler & Jenkins, 1989) indicate that even young children are capable of discovering new strategies and thus improving their problem-solving performance without direct instruction. Thus, we explored the extent to which children's performance improved over trials. We found that the greatest improvements occurred for the most difficult type of evidence—positive and hidden—and that this improvement was largely confined to the older children. Although, as noted earlier, 4- and 5-year-old children were equally likely to respond incorrectly to this type of evidence (especially on the $+-?$ pattern, with an error rate more than 70%), the 5-year-olds' performance did improve, to some extent, over repeated encounters with these types of problems. Although only about 20% of them correctly responded to the 1st positive and hidden trial, by the 3rd positive and hidden trial this increased to 40% correct, and by the 10th trial this increased to 50% correct. In contrast, about 30% of the 4-year-olds correctly responded to the 1st positive and hidden trial, and they remained at about this level throughout the experiment.

This relatively low performance might have been due to the nature of the feedback. Children never received any explicit feedback in this study. Instead, it was implicit and inconsistent as the evidence pattern for a particular problem was sequentially revealed: sometimes supporting and sometimes contradicting an earlier response. For example, an erroneous know response to $+-?$ was sometimes disconfirmed when the child opened the final box (if the problem ended as $++$) and sometimes confirmed (if it ended as $+-$). Thus, Experiment 2 was designed to include explicit and systematic feedback to explore better the extent to which children could be taught to replace the positive-capture strategy with a more advanced strategy.

Experiment 2

In Experiment 2 we adopted the second of the two commonly used methods described earlier for investigating children's performance on a particularly difficult acquisition. We used a microgenetic approach in which we provided explicit and systematic feedback on all pattern types over several days in several phases, including a pretest, two learning phases, a posttest, and a delayed transfer

test several months later. The goal was to extend the suggestive finding from Experiment 1 that 5-year-olds' performance on the positive and hidden evidence patterns appeared to improve over trials even though no explicit feedback was provided. Children were assigned to two learning conditions. In the training condition, correct and incorrect responses were pointed out and the rationale for the correct answer was provided immediately after the child responded to each question. That is, children received immediate, explicit, and consistent feedback after each box was opened. In the control condition, children received no explicit feedback from the experimenter about whether their response was correct or incorrect. (However, they could receive implicit feedback of the sort described in Experiment 1.) We predicted that 5-year-olds in the training condition would learn more effectively than their peers in the control condition, and that 4-year-olds' problem-solving performance would not benefit from the implicit feedback experience in the control condition. Experiment 2 was also designed to address strategy transfer. We designed various problems with different materials (e.g., box, stamp, and marker tasks) to explore whether and how children at different age levels transfer the learned strategy from one task to another.

Method

Participants

Twenty-eight 4- and 5-year-olds ($range = 4.4\text{--}6.4$ years) from a university laboratory preschool participated in the first four phases (pretest, Learning I, Learning II, and posttest). Children at each age level were randomly assigned to one of two conditions. The mean ages of 5-year-olds in the training ($n = 7$) and control ($n = 8$) conditions were 5.7 and 5.5, respectively, and mean ages of 4-year-olds in the training ($n = 8$) and control ($n = 5$) conditions were 4.8 and 4.9, respectively. Fifteen children were assigned to the training condition and 13 to the control condition. Of the 28 children who participated in the first four phases, 21 (eleven 4-year-olds and ten 5-year-olds) participated in the follow-up phase 7 months later (at which point the mean ages of 5-year-olds in the training and control conditions were 6.2 and 6.1, respectively, and of 4-year-olds were 5.3 and 5.5, respectively).

Materials

Three isomorphic tasks were used in this study: markers, stamps, and boxes.

Marker task. Stimulus materials were 20 marker holders with one marker inside each holder (5×3 for Day 1 or Day 4, and 5 for familiarization). Five pictures drawn in one color on white paper. A foam core rack was used to hold the marker holders upright (see Figure 5c). A piece of cardboard was used as a folding screen (for this and the other two tasks) to hide the experiment's "construction" of the target item.

Stamp task. Stimulus materials were 20 rubber stamps (5×3 for Day 1 or Day 4 problems, and 5 for familiarization). A stamp pad and five foam core racks were used, each holding three stamps and a strip of blank paper.

Box task

Stimulus materials were 35 boxes with lids (5×3 for Day 2 problems and 5×3 for Day 3 problems, and 5 for familiarization), each containing a number of the same objects that can be strung or clipped together.

The stamp task is illustrated in Figure 5. The marker task is similar to the no-context materials in Experiment 1, and the box task is similar to the task used in the second experiment in Fay & Klahr (1996).

Task and problem orders. Either stamps or markers were used on for the pretest phase on Day 1. The marker (or stamp) task consisted of 15 marker holders (or stamps), five sets each of three markers (or stamps). Fifteen boxes, five sets each of three boxes, were used for Learning I and Learning II phases on Days 2 and 3. For the posttest phase on Day 4, either stamps or markers were used, whichever the participants did not have on Day 1.

Problems, evidence patterns, and evidence types. Panel A of Table 4 shows the five problems given to participants on each day, with the evidence patterns

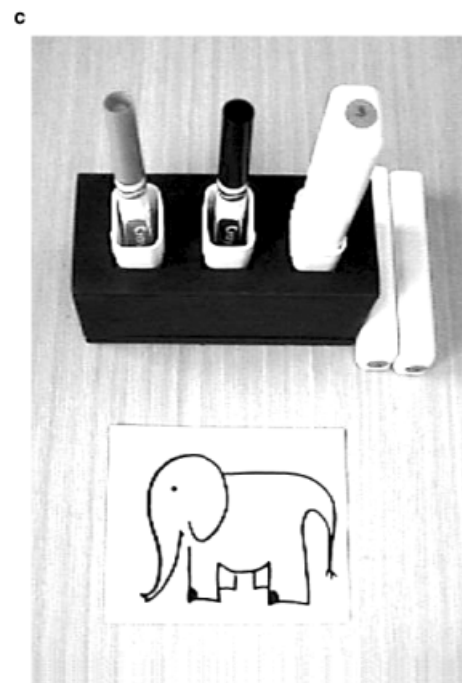
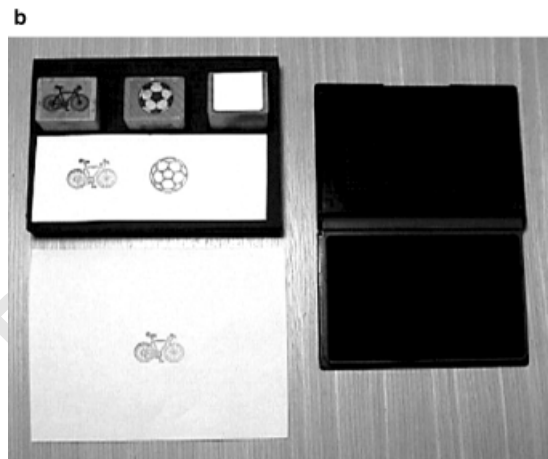


Figure 5. Examples of materials used in Experiment 2: (a) *Boxes problem.* The goal is to determine from which of the three boxes came the materials used to construct the necklace in front of the boxes. In this example, all boxes are opened, revealing a $- - +$ evidence pattern: Only the right-most box could have been used. (b) *Stamps problem.* The goal is to determine which of the three boxes contains the stamp that was used to make the picture shown on the large sheet at the bottom of the figure. The contents of each box are indicated both by the design on the box lid and the stamped pattern in front of each box. All boxes start out with the design on the box lid obscured by a piece of paper. In this example, the contents of the first two boxes, but not the third, have been revealed, yielding a $+ - ?$ pattern. (c) *Marker problem.* Two of the three markers are exposed. The goal is to determine which of the three markers was used to draw the picture. In this example, the first open box contains a nonmatching marker, the second open box contains a matching marker, and the third box has not yet been opened (corresponding to a $- + ?$ pattern of evidence; see text for full explanation).

that occur in each problem and Panel B shows the categorization of evidence patterns into various types and the frequency with which each type occurred over the full set of problems. All participants received the same fixed problem order, shown in Panel A. On each day, three of the five problems were indeterminate and the other two were determinate.

Procedure

Overview. Each of the first four phases lasted 15 to 30 min and occurred about a week apart. Participants were assigned to one of two conditions: training or control. Children's pretest performance on Day 1 was used to ensure equivalent initial performance levels in both conditions. Children were rank ordered by their pretest performance on the positive and hidden trials, the highest performers were eliminated, and the remaining ranked children were assigned alternately to each condition. Children in the control condition were never given explicit feedback about their answers, whereas children in the training condition were given explicit feedback about their responses on Days 2 and 3.

Children were tested individually in a room in their school. Two separate phases were presented: familiarization and test problems. In both phases, the items were first presented closed and then opened one at a time by the experimenter and child. The five problems shown in Panel A of Table 4 were presented in each of the four experimental phases. In the pretest, Learning I, and posttest phases, the experimental problems were preceded by a familiarization phase on the new materials (i.e., because of the change from, say, boxes to stamps). However, because the materials in the Learning II phase were the same as in the Learning I phase, no familiarization was necessary. The pretest and posttest phases were the same for both control and training conditions, whereas the procedures for the Learning I and Learning II phases were different as outlined in the following discussion and in Table 5. The procedure described is for the boxes task, but the procedure is the same for the stamps and marker holder tasks, as well.

Familiarization phase. During familiarization, the experimenter showed the child five boxes and opened each one. Two boxes contained green cubes,

Table 4

Types of Problems, Patterns of Evidence, and Their Frequencies of Occurrences (Experiment 2)

Panel A. Problem types and evidence patterns in order used in all phases					
Problem type and presentation order					
	Indeterminate	Determinate	Indeterminate	Indeterminate	Determinate
Evidence patterns as boxes are opened and contents revealed	???	???	???	???	???
	+??	+??	-??	+??	-??
	+ - ?	+ - ?	- + ?	++?	- - ?
	+ - +	+ - -	- + +	++ -	- - +
Panel B. Frequency of different types of evidence patterns					
Evidence pattern	Correct response	Pattern frequency	Evidence type	Type frequency	
???	Guess	5	Hidden	5	
-??	Guess	2	Negative and hidden	3	
- - ?	Know ^a	1			
+??	Guess	3	Positive and hidden	6	
+ - ?	Guess	2			
- + ?	Guess	1			
++?	Guess	1	Positive and visible	6	
++ -	Guess	1			
+ - +	Guess	1			
- + +	Guess	1			
- - +	Know	1			
+ - -	Know	1			

Table 5
Procedure for Experiment 2

	Phase (Timing)				
	Pretest (Day 1)	Learning I (Day 2)	Learning II (Day 3)	Posttest (Day 4)	Follow-up (7 months later)
Materials	Markers or stamps	Boxes ^a	Boxes	Stamps or markers	Markers in boxes on small tables ^b
Problems	5 boxes to familiarize 5 3-box problems	5 boxes to familiarize 5 3-box problems	5 3-box problems	5 boxes to familiarize 5 3-box problems	5 boxes to familiarize 5 3-box problems
Explicit feedback?	No	Training condition —Yes Control condition —No	Training condition —Yes Control condition —No	No	No

^aBoxes contained materials that could be joined together (e.g., plastic building blocks) or strung on a string (e.g., beads or tubular pasta) to form the target objects.

^bMaterials and procedure for the follow-up phase were similar to those used in the neutral condition of Experiment 1.

two contained yellow cubes, and one contained red cubes, so that children could see that each box held only one color of cubes and that it was possible for two boxes to have the same color and shape of object. The experimenter asked the children to point out the boxes that contained the same color objects. The boxes were then closed and put away.

Experimental phase. After the familiarization phase, the children were told they were going to play a game with boxes that have new things inside: "So here we have three boxes. You can see that they all look the same when they're closed. I'm going to open one box, and I'm going to make something using only the things from this one box. I'm going to put this screen up so that you can't see which box I'm using."

The experimenter placed a screen in front of boxes and produced a premade target, then removed the screen to reveal the target and all three closed boxes: "Now you get to see if you can figure out which box I used to make this [target]." Before any boxes were opened, and after each box was opened, children were asked the test question: "Do you know for sure which box I used to make that [target], or do you have to guess?" As in Experiment 1, the order of "know for sure" and "have to guess" were randomized throughout the experiment. Children were then asked to explain their response via one of two probe questions: "How do you know for sure?" or "Why do you have to guess?" The experimenter then opened the boxes sequentially, asking the test and probe questions after each box was opened.

Control group. The procedure for control condition was as described earlier. At no point during the any of the phases did control condition participants

receive any feedback from the experimenter about their know or guess responses or their explanations. Recall, however, that the problems themselves provide implicit feedback because the opening of subsequent boxes in a problem can either validate or invalidate a previous answer about that problem.

Training group. Children in the training group were tested as described earlier for the pretest and posttest days, with the addition of feedback in Learning I and Learning II on Days 2 and 3. In the Learning I phase, all participants in the training condition were given full feedback. In the Learning II phase, the feedback was limited to confirming the child's answer, if the child gave both a correct know or guess response and a valid explanation. Incorrect know and guess responses or invalid explanations triggered the same corrective feedback as in Learning I.

After answering the test and explanation questions the children were given feedback on their answers. Depending on their answer they were told: "You said you know for sure (or have to guess), and that's (not) right." Then the reason for the correct guess or know answer was explained by the experimenter. The explanation focused on the possibility that any remaining closed boxes could have the target item inside them, whether or not there was already a box opened with the target item inside. For example, in the situation in which one box was open and the contents matched the target, the feedback was as follows: "This is a time when you have to guess. Even though you see a box with [target] in it, there could be [targets] in these boxes also (point to closed boxes), and I could have used one of these boxes when I made this [target] (point to target)."

Procedure for follow-up session. The follow-up phase took place approximately 7 months after the four main phases of the study. Two problem orders of five problems each were used. The task was similar to the markers task used earlier, except that the experimenter and participants traveled throughout the room to premade pictures set on little tables in front of the marker boxes, similar to the neutral condition of Experiment 1.

Results

Three major dependent variables were measured. The first involves children's response to the know-guess question ("Do you know for sure which box I used to make the [target], or do you have to guess?") on each trial. For indeterminate patterns, the correct response is guess, whereas for determinate patterns, the correct answer is know. Responses were scored as 1 (*correct*) or 0 (*incorrect*).

The second measure is based on children's explanations for their answers to the question about whether they knew or had to guess. Explanations were scored as 1 (*good*) or 0 (*irrelevant*). A good explanation is both supportive of their response (saying you have to guess "because two markers match," or that you know for sure "because only one matches") and global in that the child takes more than one box (or marker or stamp) into account (saying you have to guess because "the other boxes are not yet opened"). Two observers coded approximately 25% of the trials independently, yielding an interrater reliability of more than 90%.

The third measure is the number (and proportion) of children who used an expert rule for each phase. To be categorized as an expert rule user in a particular phase, the child had to respond correctly on at least 16 of 20 trials, and on at least 4 out of the 6 positive and hidden trials.

Data were analyzed to address the following questions: (a) Does performance depend on evidence type? (b) Do older children outperform younger children? (c) Does performance improve over phase? (d) Do older children and children in the training condition improve more than younger children and children in the control condition? The results are organized in three sections: (a) overall performance, (b) performance on positive and hidden evidence patterns, and (c) children's use of the expert rule.

Overall Performance

Pretest. We first examined children's performance on each type of evidence in pretest to determine (a)

whether performance patterns were similar to those obtained in previous studies and (b) whether children's initial performance was comparable between the two conditions. Overall, children responded correctly to the hidden, negative and hidden, and positive and visible evidence types on nearly 70% of the pretest trials (see Figure 6). However, fewer than 20% of their responses were correct for positive and hidden evidence type. A 2 (age) \times 2 (condition) \times 4 (evidence type) ANOVA with evidence type as a repeated measure yielded significant differences among types, $F(3, 69) = 43.85$, $p < .0001$, but no main effects for age or condition, nor any interactions among age, condition, and evidence type. Moreover, the patterns of initial performance levels on different problem types were similar in both conditions and for both age groups, and were similar to those obtained in previous studies.

Children's overall learning from experience: mean scores at each phase. Figure 7 shows the percentage of correct responses at both age levels and in both conditions over phase. To examine 4- and 5-year-olds' correct responses to all types of problems in the training and control conditions over phase, we performed a 2 (condition) \times 2 (age) \times 5 (phase: pretest, Learning I, Learning II, posttest, and follow-up) ANOVA with phase as a repeated measure. Note that only the scores of children who participated in all five phases are included. This analysis revealed significant main effects for condition, $F(1, 68) = 11.70$, $p < .01$; for age, $F(1, 68) = 13.95$, $p < .01$; and for phase, $F(4, 68) = 10.60$, $p < .001$. This analysis

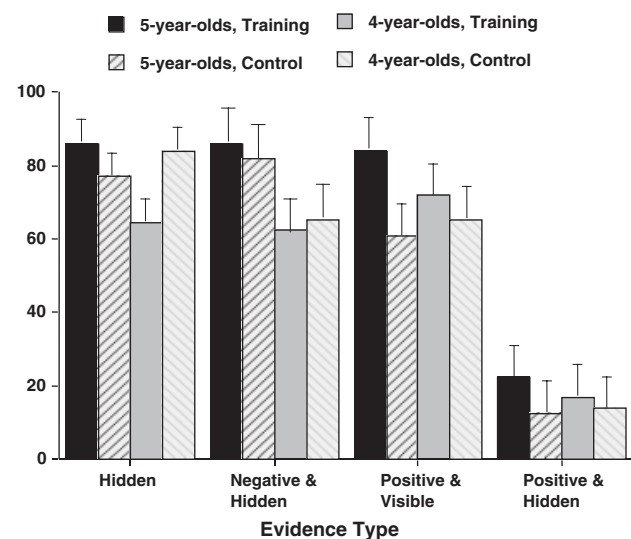


Figure 6. Percentage of correct responses to each type of evidence in pretest, with standard error bars (Experiment 2).

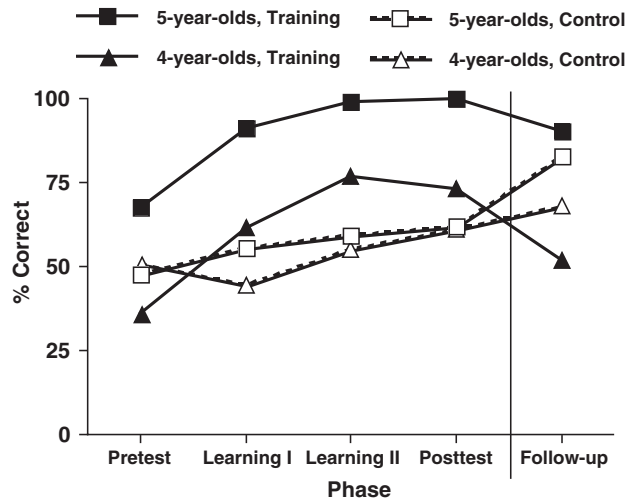


Figure 7. Percentage of overall correct responses over phase by conditions and age (Experiment 2). Vertical line indicates 7-month delay between first four phases and the follow-up phase.

also yielded a significant interaction between condition and age, $F(1, 68) = 6.62, p < .05$, and an interaction between condition and phase, $F(4, 68) = 6.50, p < .001$. To examine the performance improvement at different ages in the two conditions, a one-way ANOVA with phase as a repeated measure was performed for each age group in each condition. Overall, children improved their performance over phase but at different rates and to different extents at different ages in different conditions: 5-year-olds in the control condition, $F(4, 20) = 14.04, p < .05$; 4-year-olds in the control condition, $F(4, 16) = 2.89, ns$; 5-year-olds in the training condition, $F(4, 12) = 20.6, p < .0001$; and 4-year-olds in the training condition, $F(4, 20) = 5.96, p < .01$. Children in the training condition significantly improved their score from pretest to Learning I ($ps < .05$ for both 4- and 5-year-olds), whereas children in the control condition did not improve their performance until the final (follow-up) phase.

Performance on Positive and Hidden Evidence Type

Mean score on positive and hidden patterns at each phase. Figure 8 shows the time course of correct responses to the positive and hidden patterns by age and condition. (As shown in Table 4, in each phase there were six positive and hidden patterns distributed as follows: three + ? ? evidence patterns, two + - ? evidence patterns, and one - + ? evidence pattern). To examine 4- and 5-year-olds' performance in the training and control conditions over phase, we performed a 2 (condition) \times 2 (age)

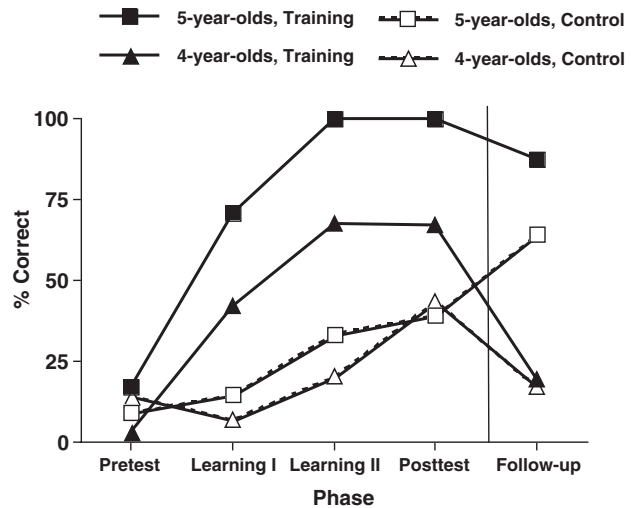


Figure 8. Percentage of correct responses to positive and hidden evidence patterns over phase by conditions and age (Experiment 2). Vertical line indicates 7-month delay between first four phases and the follow-up phase.

\times 5 (phase) ANOVA with phase as a repeated measure. (The scores of children who did not participate in the follow-up test are not included.) This analysis revealed significant main effects for condition, $F(1, 68) = 12.66, p < .005$; for age, $F(1, 68) = 7.18, p < .05$; and for phase, $F(4, 68) = 15.24, p < .0001$, and significant interactions between condition and phase, $F(4, 68) = 5.50, p < .001$, and between age and phase, $F(4, 68) = 3.74, p < .01$.

Given that there are interactions between phase and condition, and between phase and age, separate 2 (condition) \times 2 (age) ANOVAs were performed on performance on each phase. The analyses reveal no significant condition or age differences in pretest performance. On Learning I phase, children in the training condition outperformed their peers in the control condition, $F(1, 24) = 18.89, p < .001$. The main effect of age is also marginally significant, $F(1, 24) = 2.98, p = .097$. This pattern remains on Learning II phase: The main effect of condition is significant, $F(1, 24) = 15.98, p < .001$, and the main effect for age is marginally significant, $F(1, 24) = 3.68, p < .067$. By posttest phase, the main effect of condition remains significant, $F(1, 24) = 8.66, p < .01$, but there is no main effect for age. By the follow-up phase, the only significant main effect is age, $F(1, 24) = 25.69, p < .0001$.

In summary, this analysis of learning on the positive and hidden patterns indicates that: (a) children at both ages and in both conditions started with similarly low performance; (b) children in the training condition improved much more than their peers in the control condition; (c) older children

improved more than younger children throughout the learning phases; (d) by the follow-up phase 7 months later, 4-year-olds' performance decreased almost to pretest levels whereas 5-year-olds' performance remained high. These results suggest that for 5-year-olds, both indirect feedback and direct instruction were effective in facilitating transfer performance with a long delay, whereas for 4-year-olds, the short-term effectiveness of direct instruction was not long lasting.

Percentage correct in both response and explanation for + ? ? evidence patterns. Of the three positive and hidden evidence types, children encountered the + ? ? pattern most often (three trials in each phase). Therefore, we performed a detailed analysis of children's trial-by-trial performance on this evidence pattern, in this case using a performance criterion that required both a correct response and a correct explanation for that response. Figure 9 shows the percentage of children who met this criterion on each of the + ? ? patterns. A 2 (condition) \times 2 (age) \times 5 (phase) ANOVA with phase as a repeated measure was performed on percentage correct (with the performance scores combined over the three trials on each phase). Again, the scores of children who participated in the early phases but did not participate in the follow-up test are not included. This analysis revealed significant main effects for condition, $F(1, 44) = 19.60$, $p < .001$; for age, $F(1,$

44) = 15.06, $p < .005$; and for phase, $F(4, 44) = 9.39$, $p < .0001$; as well as significant interactions between condition and phase, $F(4, 44) = 5.48$, $p < .005$ and between age and phase, $F(4, 44) = 2.99$, $p < .05$.

Separate one-way ANOVAs with phase as within-subject variable were performed in each condition at each age level. The analyses reveal that 5-year-olds in the training condition and in the control condition significantly improved their performance over phase, $F(4, 8) = 10.26$, $p < .005$, and $F(4, 8) = 3.92$, $p < .05$, respectively. Furthermore, 5-year-olds in the training condition increased their scores from pretest to the learning phases, whereas their peers in the control condition did not improve their scores until the follow-up phase. In contrast, 4-year-olds in both conditions did not significantly increase their scores.

Expert Rule Use

Figure 10 shows the percentage of children who used the expert rule in each of the five phases (including the follow-up phase). This analysis is based on the 21 children who participated in all five initial phases, including the follow-up phase 7 months later (training: four 5-year-olds and six 4-year-olds; control: six 5-year-olds and five 4-year-olds). Note that the effects reported here were sufficiently strong that even with these small numbers they were statistically significant. (An analysis based on the full set of 28 children who participated in only the first four phases yielded

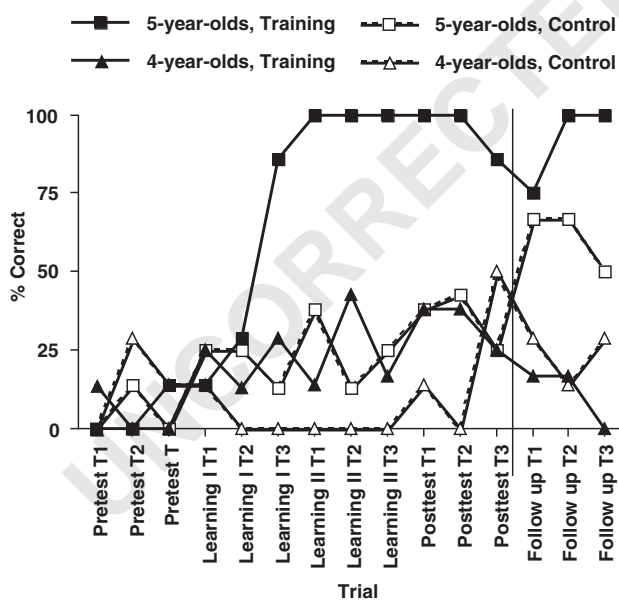


Figure 9. Percentage of children who were correct in both response and explanation on each occurrence of a + ? ? pattern (Experiment 2). Vertical line indicates 7-month delay between first four phases and the follow-up phase.

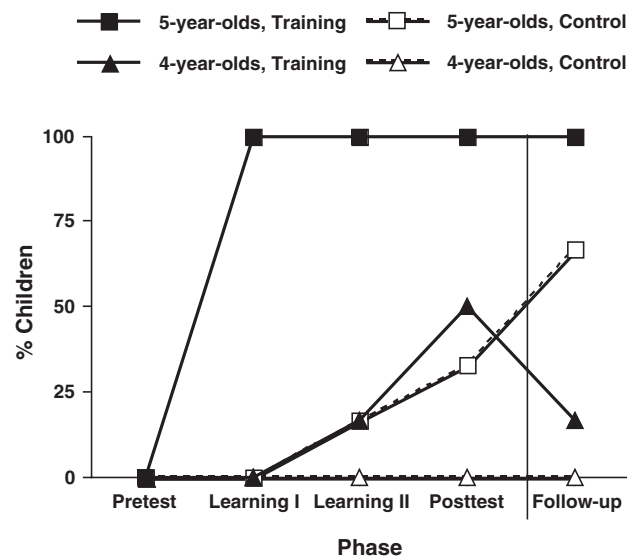


Figure 10. Percentage of children achieving an expert rule in each phase (Experiment 2). Vertical line indicates 7-month delay between first four phases and the follow-up phase.

similar results.) No 4-year-old in the control condition learned to use the expert rule on any phase. In contrast, 5-year-olds in the training condition quickly learned the expert rule. By the posttest phase, 7 of the 10 children in the training condition but only 2 of the 11 controls were using the expert rule (Fisher exact $p = .03$). However, by follow-up the training effect had been replaced by a simple age effect: Eight of the ten 5-year-olds used the expert rule, but only one of the eleven 4-year-olds did (Fisher exact $p = .002$). The performance pattern involved in the expert rule use is highly consistent with that involved in percentage of correct responses over phase.

To examine individual differences in the acquisition of the expert rule, children's learning patterns were categorized into four groups: (a) fast learn: used the expert rule in the learning phases; (b) slow learn: used the rule in posttest, the follow-up phase, or both; (c) learn and lost: used the rule in the learning phases but did not use it in posttest and follow-up phases; and (d) no learn: never used the rule. Figure 11 shows that all 5-year-olds in the training condition are in the fast learner category, and all 4-year-olds in the control condition are in the no learn category. Five-year-olds in the control condition and 4-year-olds in the training condition

fall in between, although more 5-year-olds in the control condition than 4-year-olds in the training condition learned the rule. These results are consistent with children's performance on other measures.

Discussion

The results showed that both 4- and 5-year-olds initially experienced difficulty in solving the positive and hidden patterns but improved their performance after receiving feedback. Children were assigned to each condition after consistently making errors on the positive and hidden patterns, and children at both ages were equally likely to use the positive-capture strategy. However, children at different ages responded differently to the repeated problem-solving experience in the control condition and to the explicit feedback in the training condition. As noted earlier, even in the control condition, problem-solving experience provided subtle, implicit feedback. Five-year-olds proved more sensitive to such implicit feedback and learned from it, whereas 4-year-olds failed to do so. Developmental differences in the training condition are evident in several aspects: Older children learned more readily from feedback, improved their performance to a greater extent, and transferred the learned correct strategy

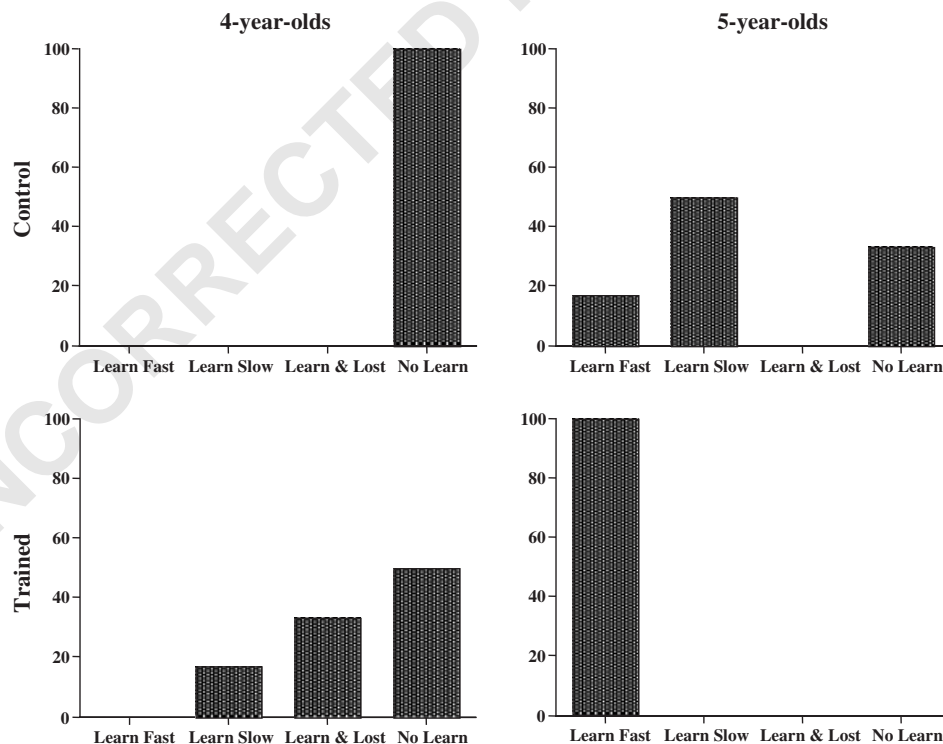


Figure 11. Percentage of children following various learning patterns in Experiment 2.

across tasks (i.e., from boxes to markers, etc.) more effectively.

General Discussion

Children's poor performance on positive and hidden patterns and their use of the positive-capture strategy could stem from several sources. One source might be the lack of socially meaningful consequence in earlier versions of the task (e.g., Fay & Klahr, 1996). Thus, we conjectured that in situations in which a judgment of determinacy or indeterminacy would result in credit or blame being assigned, children might consider the evidence presented to them more carefully. Another source of positive-capture errors might be children's lack of extensive problem-solving experience and explicit feedback. That is, these types of problems are novel, and it might take several exposures to them before children fully understand how to map these evidence patterns to know or guess responses. In this article we report on our attempts to address both of these issues. Our results revealed that children continued to use the positive-capture strategy in solving the indeterminacy problems under various situations. Presenting problems in socially richer contexts—at least as operationalized here—had no effect on children's performance (Experiment 1). Moreover, implicit feedback inherent in simply experiencing a sequence of problems had little effect on 4-year-olds' performance. On the other hand, explicit feedback facilitated both 4- and 5-year-olds' performance, and was particularly effective for older children (Experiment 2). These results are discussed in the following sections.

Robustness of Positive Capture

Children in the present research experienced difficulty in solving particular types of indeterminate patterns but performed reasonably well in solving determinate patterns. This finding is consistent with previous studies (Byrnes & Overton, 1986; Pieraut-Le Bonniec, 1980; Scholnick & Wing, 1988) of 7- to 10-year-olds solving more complex problems. The present experiments—using tasks with various forms, contents, and contexts—replicated Fay and Klahr's (1996) results that children's understanding of indeterminacy is limited: They mainly use a positive-capture strategy that correctly evaluates the determinacy or indeterminacy of all but the positive and hidden evidence patterns.

One of the aims of the present research was to explore whether young children's performance would improve in more supportive test contexts.

We reasoned that their use of the positive-capture strategy in previous studies might have been due to the relatively abstract nature of the tasks, and thus their performance might not reflect their competence. If so, testing children in a more socially meaningful context might motivate them to bring relevant knowledge to bear.

Contrary to our initial conjecture, children performed equally well (or poorly) in the neutral and context conditions and, within the context condition, on credit and blame problems. One could interpret this as evidence that the positive-capture strategy is so deeply entrenched in young children—both 4- and 5-year-olds—that they use it regardless of task context (neutral and social contexts), problem format (three- and four-box problems), and problem content (box, marker, and stamp tasks). However, as noted at the outset, there are two potential problems with our operationalization of the social context. One is that the cover story simply might not have been sufficiently salient for children to be motivated to resolve the assignment of credit or blame. Another is that current theories of how young children prefer to assign credit and blame do not make an unambiguous prediction about the direction of the effect we were hoping to induce.

Nevertheless, Experiment 1 did provide a strong replication of earlier investigations of positive capture, even with substantial changes in the particular materials and cover stories. At the least, it provided further evidence that the positive-capture strategy is a genuine and general phenomenon. Yet, even so, such a deep misconception could be overcome. More specifically, extensive and specific problem-solving experience and especially explicit feedback proved effective in enhancing the acquisition of a correct strategy and encouraging abandonment of the positive-capture strategy.

Developmental Differences in Learning a New Strategy From Experience and Training

Although Experiment 1 was not designed as a microgenetic study to examine learning, its results suggested that even without instruction, young children could learn from their problem-solving experience. However, even this limited learning was confined mainly to the 5-year-olds. In Experiment 2, we used a microgenetic design to explore the effect of direct, explicit training on children's ability to replace the positive-capture strategy with a correct strategy.

In Experiment 1 and in the control condition of Experiment 2, the sequential opening procedure

afforded implicit feedback. At least the 5-year-olds might have noticed that their earlier premature determinate response to an evidence pattern was in fact incorrect as the boxes were opened. Yet, this type of feedback was not very informative for several reasons. First, children might not notice the discrepancy between their responses and the subsequent outcomes. Second, the implicit feedback was not always consistent. For example, on a determinate problem, after a child incorrectly responded "know" to the $- + ?$ pattern, the mistaken attribution of determinacy would be reinforced positively, just a moment later, when the last box was opened, revealing the $- + -$ pattern (and thus the incorrect response was not disconfirmed). In contrast, on an indeterminate problem, that same incorrect response to the $- + ?$ pattern would be contradicted when the last box was opened, revealing the $- + +$ pattern. Younger children were not expected to learn from such implicit, inconsistent, nonimmediate and indirect feedback. However, in the training condition, after each trial, the child received information from the experimenter concerning whether his or her response was correct or incorrect, as well as the rationale behind the correct answer.

Our results show that even 5-year-olds are capable of mastering this difficult reasoning strategy if they are provided with explicit feedback. To a lesser extent, 5-year-olds' performance also benefits from extensive problem-solving experience with implicit feedback. Moreover, they are able to transfer that learned strategy to different contexts even after a 7-month delay. These findings extend prior research demonstrating that elementary school children can overcome what appear to be stubborn misconceptions when they receive explicit instruction and feedback (Chen & Klahr, 1999; Klahr & Carver, 1988; Siegler, 1996; Siegler & Chen, 1998).

There were several interesting differences in the performance of the 4- and 5-year-old groups in the training condition. The first difference is the magnitude of the learning. Five-year-olds in the training condition reached nearly perfect performance for the later trials in both their responses and explanations, and almost all of them used the expert rule in solving the problems in later phases. In contrast, 4-year-olds in the same condition never reached more than 50% correct. Nevertheless, nearly half of them used the expert rule in the posttest phase.

The second aspect of age difference is the rate of learning. Five-year-olds managed to abandon the positive-capture strategy after they received feedback on the first and second $+ ? ?$ evidence patterns. By the third $+ ? ?$ evidence pattern in Learning I,

more than 80% of the 5-year-olds gave correct responses and explanations, and all of them did so when they subsequently encountered the first $+ ? ?$ pattern in Learning II. Also, more than 80% of the 5-year-olds in the training condition achieved the expert rule by Learning I, and 100% of them used the rule in later phases. In contrast, 4-year-olds in the training condition did not significantly improve their performance in both responses and explanations until the last posttest trial.

Of particular interest is the fact that although both older and younger children showed no differences in their initial scores on the positive and hidden types of evidence, there was an age effect when initial responses to all evidence types were combined. This suggests that the older children have become proficient at dealing with the general evidence-evaluation process when it is instantiated by patterns that are clearly determinate or indeterminate. And although they still find the positive and hidden types as difficult as do the younger children, they are likely to devote fewer processing resources to the general cognitive and attentional demands of this type of problem and are thus better able to learn from both explicit and implicit feedback with repeated exposure to indeterminacy problems.

The third aspect of learning involves the generalization of the acquired expert rule. All 5-year-olds who acquired the expert rule maintained its use from the learning phases to the posttest phase to the follow-up phase. They showed the ability to generalize the acquired rule even after a 7-month delay, whereas 4-year-olds experienced difficulty in generalization when the gap between the learning situation and the new problems became more distant. Thus, feedback proved effective for 5-year-olds but had no lasting effect on younger children. Similarly, problem-solving experience alone proved effective in facilitating 5-year-olds' reasoning but failed to enhance younger children's performance. It is worth noting that the follow-up phase represents a rigorous test of young children's ability to transfer the learned strategy to a new task that differed from the earlier tasks in superficial features with long delay. Most previous studies on analogical transfer with children and adults involve presenting training problems and asking participants to solve analogous tasks minutes, hours, or days later, but manipulation of the time interval beyond a few months is rare (see Barnett & Ceci, 2002). The present results demonstrated 5-year-olds' impressive ability to transfer the acquired strategy with a long delay.

Implications for Early Science Instruction

As indicated by Experiment 1, and earlier work on similar problems (e.g., Fay & Klahr, 1996), most preschoolers were poorly prepared to deal with certain types of evidence patterns, and they were likely to characterize incorrectly indeterminate evidence as determinate. However, the results of Experiment 2 showed that although they started from similar position with respect to the hardest problems, the 5-year-olds were much more responsive to feedback, both direct and indirect, than were the 4-year-olds. This suggests that because the 5-year-olds had already mastered the easier patterns and general procedure associated with reasoning about determinate and indeterminate evidence, they were more able to assimilate the feedback and more rapidly and effectively replace a flawed strategy with a correct strategy.

Our results are consistent with earlier findings with older children showing that both indirect feedback and direct instruction play a role in overcoming young children's misconceptions (e.g., Chen & Klahr, 1999; DeLoache, Miller, & Pierroutsaks, 1998; Gentner & Gentner, 1983; Klahr & Carver, 1988; Siegler, 1976, 1981, Siegler & Chen, 1998). Although several of these studies have demonstrated that older children can learn—via discovery or instruction—more complex principles and rules, the present research demonstrates that with structured experience and instruction even 5-year-olds can master the fundamental concept of indeterminacy.

We believe that this process was facilitated by presenting several analogous tasks, all having different surface contexts but the same underlying logical structure. This enabled children to construct a generalized schema that they could subsequently transfer to novel tasks (c.f. Brown, Kane, & Echols, 1986; Butterfield, Slocum, & Nelson, 1993; Chen & Daehler, 1989; Gentner, 1983, 1989; Gick & Holyoak, 1983, 1987). Thus, one approach to early science instruction would be to present multiple analogous tasks that share the same logical structure so that young children acquire the underlying principles or concepts from the experience in working on these tasks and from the explicit, direct feedback. This suggests a reconsideration of the current tendency in much of early science education to favor rich and varied content, rather than a consistent underlying structure. Such a focus on the logical content of science instruction would enable researchers and educators to determine the cognitive demands of the full range of substantive topics, both domain-specific knowledge and domain-general processes.

Conclusions

The present research reveals that the positive-capture strategy is a robust phenomenon in young children's reasoning. Both 4- and 5-year-olds made positive-capture errors in both neutral and social contexts and in solving problems with different formats (three- and four-box problems) and contents (box, marker, and stamp tasks). Yet, it was also evident that 5-year-olds learned from both problem-solving experience and explicit feedback, and 4-year-olds' performance on indeterminacy problems also benefited from explicit feedback but not from implicit feedback. Older children's learning proved more effective in magnitude, rate, and generalization. Although most 5-year-olds initially used the positive-capture strategy in solving indeterminacy patterns, they learned readily from feedback to recognize and correctly explain both determinate and indeterminate situations, and the acquired strategy was effectively generalized across different tasks. The present findings point to the feasibility and importance of early science instruction on young children's acquisition of scientific concepts and strategies.

References

- Acredolo, C., & Horobin, K. (1987). The child's relational reasoning and tendency toward premature closure. *Developmental Psychology*, 23, 13–21.
- Barnett, S., & Ceci, S. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, 128, 612–637.
- Bjorklund, D. F., & Green, B. L. (1992). The adaptive nature of cognitive immaturity. *American Psychologist*, 47, 46–54.
- Braine, M. D. S., & Rumin, B. (1983). Logical reasoning. In J. H. Flavell & E. M. Markman (Eds.) P. H. Mussen (Series Ed.), *Handbook of child psychology: Vol. 3. Cognitive development* (pp. 263–340). New York: Wiley.
- Brainerd, C. J., & Allen, T. W. (1971). Training and generalization of density conservation: Effects of feedback and consecutive similar stimuli. *Child Development*, 42, 693–704.
- Brainerd, C. J., & Brainerd, S. H. (1972). Order of acquisition of number and quantity conservation. *Child Development*, 43, 1401–1406.
- Brown, A. L., Kane, M. J., & Echols, C. H. (1986). Young children's mental models determine analogical transfer across problems with a common goal structure. *Cognitive Development*, 1, 103–121.
- Butterfield, E. C., Slocum, T. A., & Nelson, G. D. (1993). Cognitive and behavioral analyses of teaching and transfer: Are they different? In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 192–257). Stamford, CT: Ablex.
- Byrnes, J. P., & Beilin, H. (1991). The cognitive basis of uncertainty. *Human Development*, 34, 189–203.

- Byrnes, J. P., & Overton, W. F. (1986). Reasoning about certainty and uncertainty in concrete, causal, and propositional contexts. *Developmental Psychology*, 22, 793–799.
- Chen, Z., & Daehler, M. W. (1989). Positive and negative transfer in analogical problem-solving by 6-year-olds. *Cognitive Development*, 4, 327–344.
- Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development*, 70, 1098–1120.
- Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, 17, 391–416.
- DeLoache, J. S., Miller, K. F., & Pierroutsaks, S. L. (1998). Reasoning and problem solving. In W. Damon (Series Ed.) & D. Kuhn & R. S. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2: Cognition, perception & language* (5th ed., pp. 801–850). New York: Wiley.
- Dunn, J. (1988). *The beginnings of social understanding*. Cambridge, MA: Harvard University Press.
- Fabricius, W. V., Sophian, C., & Wellman, H. M. (1987). Young children's sensitivity to logical necessity in their inferential search behavior. *Child Development*, 58, 409–423.
- Falmagne, R. J., Mawby, R. A., & Pea, R. D. (1989). Linguistic and logical factors in recognition of indeterminacy. *Cognitive Development*, 4, 141–176.
- Fay, A. L., & Klahr, D. (1996). Knowing about guessing and guessing about knowing: Preschoolers' understanding of indeterminacy. *Child Development*, 67, 689–716.
- Field, D. (1987). A review of preschool conservation training: An analysis of analyses. *Developmental Review*, 7, 210–251.
- Fincham, F. D. (1985). Outcome valence and situational constraints in the responsibility attributions of children and adults. *Social Cognition*, 3, 218–283.
- Gelman, R. (1969). Conservation acquisition: A problem of learning to attend to relevant attributes. *Journal of Experimental Child Psychology*, 7, 167–187.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155–170.
- Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 199–241). Cambridge, England: Cambridge University Press.
- Gentner, D., & Gentner, D. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 199–241). Hillsdale, NJ: Erlbaum.
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogic transfer. *Cognitive Psychology*, 15, 1–38.
- Gick, M. L., & Holyoak, K. J. (1987). The cognitive basis of knowledge transfer. In S. M. Cormier & J. D. Hagman (Eds.), *Transfer of learning: Contemporary research and applications* (pp. 9–46). San Diego, CA: Academic Press.
- Hamilton, V. L., Blumenfeld, P. C., & Kushler, R. H. (1988). *Journal of Personality and Social Psychology*, 54, 34–48.
- Harris, P. L., & Nunez, M. (1996). Understanding of permission rules by preschool children. *Child Development*, 67, 1572–1591.
- Horobin, K., & Acredolo, C. (1989). The impact of probability judgments on reasoning about multiple possibilities. *Child Development*, 60, 183–200.
- Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic.
- Johnson, C. N., & Maratsos, M. P. (1977). Early comprehension of mental verbs: Think and know. *Child Development*, 48, 1743–1747.
- Johnson, C. N., & Wellman, H. M. (1980). Children's developing understanding of mental verbs: Remember, know, and guess. *Child Development*, 51, 1095–1102.
- Klahr, D. (2000). *Exploring science: The cognition and development of discovery processes*. Cambridge, MA: MIT Press.
- Klahr, D., & Carver, S. M. (1988). Cognitive objectives in a LOGO debugging curriculum: Instruction, learning, and transfer. *Cognitive Psychology*, 20, 362–404.
- Klahr, D., Fay, A. L., & Dunbar, K. (1993). Heuristics for scientific experimentation: A developmental study. *Cognitive Psychology*, 25, 111–146.
- Kuhn, D. (1995). Microgenetic study of change: What has it told us? *Psychological Science*, 6, 133–169.
- Kuhn, D., Garcia-Mila, M., Zohar, A., & Andersen, C. (1995). Strategies of knowledge acquisition. *Monographs of the Society for Research in Child Development*, 60, (4, Serial No. 245).
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108, 480–498.
- Light, P., Blaye, A., Gilly, M., & Girotto, V. (1989). Pragmatic schemas and logical reasoning in 6- to 8-year-old children. *Cognitive Development*, 4, 49–64.
- Miller, P. H., & Coyle, T. R. (1999). Developmental change: Lessons from microgenesis. In E. K. Scholnick & K. Nelson (Eds.), *Conceptual development: Piaget's legacy* (pp. 209–239). Mahwah, NJ: Erlbaum.
- Miscione, J. L., Marvin, R. S., O'Brien, R. G., & Greenberg, M. T. (1978). A developmental study of preschool children's understanding of the words "know" and "guess." *Child Development*, 49, 1107–1113.
- Moore, C., Bryant, D., & Furrow, D. (1989). Mental terms and the development of certainty. *Child Development*, 60, 167–171.
- Morris, B. J., & Sloutsky, V. (2002). Children's solutions of logical versus empirical problems: What's missing and what develops? *Cognitive Development*, 76, 1–22.
- Nelson-le Gall, S. A. (1984). Children's and adults' assignment of blame for personal injury. *The Journal of Psychology*, 117, 135–142.
- Osherson, D. N. (1975). Models of logical thinking. In R. Falmagne (Ed.), *Reasoning: Representation and process in children and adults* (pp. 81–93). Hillsdale, NJ: Erlbaum.
- Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.
- Piérault-Le Bonniec, G. (1980). *The development of modal reasoning: The genesis of necessity and possibility notions*. New York: Academic Press.
- Ruffman, T. K., & Olson, D. R. (1989). Children's ascription of knowledge to others. *Developmental Psychology*, 25, 601–606.

- Schleifer, M., Schultz, T. R., & Lefebvre-Pinard, M. (1983). Children's judgements of causality, responsibility, and punishment in cases of harm due to omission. *British Journal of Developmental Psychology*, 1, 87-97.
- Scholnick, E. K., & Wing, C. S. (1988). Knowing when you don't know: Developmental and situational considerations. *Developmental Psychology*, 24, 190-196.
- Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive Psychology*, 8, 481-520.
- Siegler, R. S. (1981). Developmental sequences within and between concepts. *Monographs of the Society for Research in Child Development*, 46, (2, Serial No. 189).
- Siegler, R. S. (1995). How does change occur: A microgenetic study of number conservation. *Cognitive Psychology*, 28, 225-273.
- Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. Oxford, England: Oxford University Press.
- Siegler, R. S., & Chen, Z. (1998). Developmental differences in rule learning: A microgenetic analysis. *Cognitive Psychology*, 36, 273-310.
- Siegler, R. S., & Crowley, K. (1991). The microgenetic method: A direct means for studying cognitive development. *American Psychologist*, 46, 606-620.
- Siegler, R. S., & Jenkins, E. A. (1989). *How children discover new strategies*. Hillsdale, NJ: Earlbaum.
- Sodian, B., Zaitchik, D., & Carey, S. (1991). Young children's differentiation of hypothetical beliefs from evidence. *Child Development*, 62, 753-766.
- Somerville, S. C., Hadkinson, B. A., & Greenberg, C. (1979). Two levels of inferential behavior in young children. *Child Development*, 50, 119-131.
- Tschirgi, J. E. (1980). Sensible reasoning: A hypothesis about hypotheses. *Child Development*, 51, 1-10.
- Wimmer, H., Hogrefe, G. J., & Perner, J. (1988). Children's understanding of informational access as a source of knowledge. *Child Development*, 59, 386-396.
- Wollman, W., Eylon, B. S., & Lawson, A. E. (1979). Acceptance of lack of closure: Is it an index of advanced reasoning? *Child Development*, 50, 656-665.

UNCORRECTED PROOF