COGNITION AND INSTRUCTION, *23*(4), 523–543

The Role of Encoding in the Self-Explanation Effect

Robert J. Calin-Jageman
Department of Biology
Georgia State University

Hilary Horn Ratner
The Graduate School
Wayne State University

We examined the relation between self-explaining and encoding among kindergartners. For 5 days, children ($n = 27$) took turns solving addition problems with an adult expert who always used an advanced addition strategy. During the game, children explained the expert's answers (Explain-Expert), explained their own answers (Explain-Novice), or did not generate explanations (Control). Encoding of the expert's strategy was measured each day by asking children to describe how the expert had solved the last problem. Explain-Expert children encoded more and learned more than children in the Control group; Explain-Novice children showed neither advantage. The Explain-Expert group also acquired the expert's strategy more rapidly and used it more frequently than the other groups. These results suggest that explanations enhance learning in part by facilitating encoding.

Requiring students to explain to themselves to-be-learned material facilitates learning (Chi, De Leeuw, Chiu, & LaVancher, 1994). This result has been dubbed the *self-explanation effect*, and it is robust across a variety of topics and learning materials. For example, self-explaining can help students learn how to better calculate compound interest (Renkl, Stark, Gruber, & Mandl, 1998), solve analytic reasoning problems (Nueman & Schwarz, 1998), and write computer programs (Bielaczyc, Pirolli, & Brown, 1995).

Requests for reprints should be sent to Robert Calin-Jageman, Department of Biology, Georgia State University, P.O. Box 4010, Atlanta, GA 30302. E-mail: rcalinjageman@gsu.edu

Two related theories have been proposed to account for the self-explanation effect (Chi, 2000). One approach, known as the *incomplete-texts* approach, suggests that self-explaining encourages students to draw inferences to fill in information omitted from their study materials (VanLehn, Jones, & Chi, 1992). Another approach, *repairing mental models*, suggests that the focus of self-explaining is not on the inadequacies of the study materials but on the deficiencies in the student's mental model of the topic (Chi, 2000). According to this view, self-explaining helps children detect and repair inaccuracies in their mental models of the topic domain.

These approaches to the self-explanation effect are distinct, but they both imply that the cognitive efforts elicited by self-explaining have an impact on students' memories. Specifically, both imply that self-explaining facilitates encoding, the initial process of storing information into memory. In the incomplete-text approach, self-explanations help students infer information omitted from their study materials. This new information is encoded into memory and then available to facilitate later performance. Students who do not self-explain do not infer the omitted information and thus encode less information while studying. This assumption was specifically built in to a computational model of the incomplete text approach (VanLehn et al., 1992). Similarly, the repairing mental models approach suggests that self-explaining focuses attention on material that contradicts or completes a student's understanding of a topic. This new material is encoded into memory through the repaired knowledge, thus improving the student's understanding of the topic. Students who do not self-explain are less likely to pay attention to this material and encode it into memory. Thus, *both* approaches predict that self-explaining facilitates encoding.

Although both accounts of the self-explanation effect predict a strong relation with encoding, there is relatively little empirical data demonstrating this relation. For instance, Chi, Bassok, Lewis, Reimann, and Glaser (1989) found that good learners who frequently self-explain refer back to the study materials less frequently than poor learners who do not frequently self-explain. VanLehn et al. (1992) interpreted this finding to indicate that good learners had better encoded their study materials and thus did not need to review them while solving problems on their own. This account is certainly plausible; however, it is not clear from the finding whether or how encoding may have contributed to the self-explanation effect. One goal of this study is to provide stronger evidence for a link between encoding and the benefits of self-explaining information.

SELF-EXPLAINING FOR YOUNG STUDENTS

A second aim for this study is to explore the significance of encoding in self-explanations among young children. Research on the self-explanation effect has only recently been extended to young students because the original protocols required participants to explain materials presented in text, and this literacy requirement seemed to preclude young learners. Siegler (1995), however, has developed a modified protocol in which students are asked to explain the actions of an adult expert. This approach eliminates reading requirements, and it has been used to show that the benefits of self-explaining extend even to very young children. For example, children prompted to explain the actions of an adult expert can improve their conservation skills (preshoolers, Siegler, 1995), acquire an advanced tic-tac-toe strategy (kindergarten–second graders; Crowley & Siegler, 1999), and enhance their theories of balance (kindergarten–third graders; Peters, Messer, Smith, & Davey, 1999; Pine & Messer, 2000).

Although these studies show that the benefits of self-explaining extend to younger children and are not limited to interactions with text, the effects are not well understood. The few studies of the self-explanation effect with young children have focused on demonstrating the effect rather than identifying the mechanisms that underlie it. Furthermore, it is unclear whether theories linking self-explaining and encoding (Chi, 2000; Neuman & Schwarz, 2000; VanLehn, Jones, & Chi, 1992) would apply to younger learners. Not only is age of the learner relevant, but also the change in the task. Self-explanation experiments conducted with young children are explicitly social, requiring interaction between two people and requiring that a novice explain another person's expert behavior rather than words within a text. The social context may lead children to enact different types of learning processes, including appropriation (Rogoff, 1990) and perspective-taking (Newman, 1986), and facilitated encoding may be less relevant.

In addition, it is not clear if enhanced encoding would benefit young children in the same way that it does older students. For example, Miller (2000) has suggested that young children may display a utilization deficiency; they may encode new information but not benefit from it due to inefficiencies in strategy use, constraints on working memory resources, and motivation. On the other hand, there is some empirical evidence suggesting that children can benefit from their memories of an adult's behavior. Specifically, Ratner, Wendorf, and Hill (2004) have found that kindergartners who participated in a series of memory games with an adult improved more if they also reproduced more of the adult's routine in the game. These considerations make it clear that the relations among self-explaining, encoding, and learning cannot necessarily be extrapolated from older to younger students.

THIS STUDY

In this experiment we attempt to directly assess the relation between self-explaining and encoding among kindergartners. As a learning task, we developed a simple addition skills game. Children could play this game on their own or take turns with an adult expert (the first author). The two-person version of the game provided an entertaining context in which to have children observe an advanced addition strategy. Specifically, the expert overtly used the *count-all* strategy, an addition strategy the children had yet to develop.

The addition game also allowed us to periodically measure children's encoding of the expert's strategy. After each round of play, children were asked to demonstrate how the expert had solved the last problem. We considered responses to this prompt to reflect children's overall encoding of the demonstrated strategy.

We designed our experimental conditions to answer two questions: (a) Does self-explaining facilitate encoding? (b) Is this facilitation necessary for improving learning? To address the first question, we varied whether or not children were prompted to explain the expert's reasoning. Children in the Control group were not prompted for explanations. Children in the Explain-Expert group were asked "How did I [the expert] know that" after each of the expert's turns. The Explain-Expert condition represents the standard self-explanation condition in which students explain either the writing or the actions of an expert. Thus, we expected children in the Explain-Expert condition to learn more than children in the Control group. The critical comparison was whether or not the Explain-Expert condition would similarly facilitate encoding.

To investigate whether or not facilitated encoding is necessary for the self-explanation effect, we attempted to dissociate the explanation process from encoding of the expert's strategy. In a third experimental condition, children in the Explain-Novice group were asked "How did you know that?" after each of their turns. This prompt requires children to generate explanations, but does not direct their attention to the expert's strategy. We thus hypothesized that the Explain-Novice condition would not facilitate encoding of the expert's strategy relative to the Control condition. If encoding of the expert's strategy is necessary for the self-explanation effect, this condition should also fail to facilitate learning.

To ensure a detailed understanding of the interrelationships among self-explaining, encoding, and strategy acquisition, we observed children over multiple sessions over a short period of time and analyzed each trial intensively. This microgenetic method is particularly useful for observing strategy acquisition (e.g., Siegler, 2000) because cognitive changes do not necessarily proceed in a linear fashion. Regressions as well as progressions; transitional states that are temporary, but necessary; and uneven generalizations all can occur and would go undetected if not analyzed across frequent episodes. This approach is also critical to evaluate Siegler's (1996) theory that multiple strategies can coexist during the same period, and that experience leads to modifications in current strategies as well as new strategies. An implication of this approach is that children playing the addition game might use several strategies but that, over the course of the study, the relative frequency of their use should change and that certain patterns of use should predict greater or lesser success in the task.

Siegler (1995) previously contrasted the benefits children accrue from generating Explain-Expert versus Explain-Novice explanations. As predicted here, he showed that children prompted to explain an expert's reasoning (Explain-Expert) learn more than children who prompted to explain their own reasoning (Explain-Novice). Siegler's study did not, however, measure encoding. Moreover, the results are difficult to interpret. Children in his Explain-Novice condition were prompted *before* receiving feedback and thus explained many incorrect answers. Children in the Explain-Expert group only explained the expert's correct answers. It is unclear, then, whether the observed differences in learning were due to explaining another's reasoning or simply explaining correct reasoning. This experiment avoids this confound by prompting children in the Explain-Novice group *after* feedback. Thus, children in the Explain-Novice group were asked either "How did you know that?" or "How could you have known that?" after each of their turns, with the prompt being determined by their accuracy on that turn.

To summarize, children played an addition game with an adult expert who repeatedly demonstrated an advanced addition strategy. During the game, children either (a) explained the expert's answers (Explain-Expert), (b) explained their own answers (Explain-Novice), or (c) did not generate explanations (Control). Encoding of the expert's strategy was measured after each round of game play. It was hypothesized that Explain-Expert children would have an advantage in learning and encoding relative to the Control group, whereas Explain-Novice children would have neither advantage.

METHOD

Participants

Participants were drawn from the kindergarten classes of two elementary schools in the same district in the Detroit metropolitan area. The district enrollment included 31% economically disadvantaged children (Michigan free-lunch program) and ranked in the 40th percentile statewide for elementary math education.

To select participants, a pretest was administered to 18 boys and 21 girls. Criteria for inclusion are discussed in the procedures section. Of the 39 children pretested, 11 (8 girls, 3 boys) exceeded the criteria for inclusion. These children did not complete any further experimental sessions. One child qualified but withdrew after the first experimental session. The remaining 27 children were randomly assigned to one of three groups: Control ($n = 9$, 6 boys, 3 girls), Explain-Novice ($n = 9$, 5 boys, 4 girls), and Explain-Expert ($n = 9$, 4 boys, 5 girls). Children were an average age of 5 years and 3 months (range = 4 years, 11 months to 5 years, 11 months). Children completed the experiment early in their first semester of kindergarten, prior to any classroom training in addition skills.

Materials

Problems. The addition problems for the experiment were two-addend problems with positive integer addends greater than 1 and sums between 6 and 10 (inclusive). To focus on children's counting strategies, "double" problems (both addends of the same value) were not included. Previous research has shown that elementary-school children solve double problems with retrieval strategies more often than nondouble problems (Ashcraft, 1992). To minimize differences in difficulty in the problem set, all problems were presented with the larger addend first. Addition problems in which the larger addend comes second are more difficult for young children than problems in which the larger addend comes first (Geary, Brown, & Samaranayake, 1991).

These criteria allowed for 11 unique addition problems. The problem "6 + 4" was dropped, to leave 10 problems. These were split into two sets, A and B, each having five problems. Test A served as the pretest for half the children; Test B for the remainder. During each experimental session, children solved four problems (randomly drawn without replacement from set A or B) and observed the expert (first author) solving four problems (randomly drawn without replacement from the opposite set). For each child, the observed–solved sets were rotated through the experimental sessions.

Math game. Problems were presented via a board game called the "The Treasure-Chest Game." The game consisted of a laminated picture of a treasure chest with a two-part "lid"; each side of the lid could be opened and closed independently. A handle in the middle of the treasure chest, however, could open both sides of the lid at once.

For each problem, the treasure chest was completely opened. The expert presented the first addend in chips and asked the child to identify the quantity. The child then put the first addend in the treasure chest by moving the chips onto the chest and closing the left side of the lid over them. The second addend was then presented and quantified. The expert placed the second addend in the chest by moving them onto the chest and closing the right side of the lid over them. The task was then to determine the total number of chips in the treasure chest. Answers could be checked by opening the chest and counting the actual number of chips placed inside.

Children played the treasure-chest game alone during the pretest; they took turns with the expert (first author) during the experimental sessions. During their turns, children were given the following prompt: "You put in X chips and then I put in Y more. How many are in the whole treasure chest?" Once an answer was given, the lid was lifted and the child identified the correct quantity. Correct answers were praised by saying "You got it right, good job." For incorrect answers, the expert said, "Oh it's Z, so X + Y is Z."

On the expert's turns, the experimenter would say, "You put in X chips and then I put in Y more. Now I have to figure out how many are in the whole treasure chest." The expert always solved problems by overtly using a "count-all" strategy (counting each addend on fingers separately, and then counting their sum).

Encoding task. When playing the treasure-chest game, the expert used an overt counting strategy. To measure which aspects of that strategy children encoded, they were asked to reproduce the expert's behavior. Specifically, children were given the following prompt at the end of each session:

> When we played the treasure-chest game I did the last one. You put in X pieces of treasure and then I put in Y pieces. It was my turn to figure out how many pieces of treasure were in the whole treasure chest. What did I do to figure that one out? Show me what I said and what I did.

Children often responded verbally to this prompt (e.g., "you counted"). Children were then encouraged to actually demonstrate the expert's strategies (e.g., "Show me what I did"). Children were not given feedback after making their responses.

Video equipment. All portions of the experiment were videotaped with standard video recording equipment.

Procedure

Pretest. Children were introduced to the treasure-chest game and solved a practice problem (2 + 1). Children were then asked to solve five problems (set A or set B). The pretest ended with the expert asking to take a turn and solving a problem. After this demonstration, children's encoding of the expert's strategy was measured (see earlier).

Children were selected for inclusion on the basis of their pretest performance. Children were excluded if they (a) could not accurately identify chip quantities corresponding to each addend, (b) used a count-all strategy, or (c) scored better than 50% on the pretest. All children met the first criterion; 11 jointly exceeded the second and third criteria.

Experimental sessions. In each session, children took turns with the expert playing the treasure-chest game, solving four problems and observing the expert solve four problems. Interleaving children's observations of the expert strategy with measurements of their performance helped maintain children's interest in the experiment and eliminated the need for separate training and measurement sessions.

Each child completed four sessions over the course of 1 week. During the treasure chest game children were prompted for explanations. Children in the Explain-Expert group were asked "How did I [the expert] know that?" after each of the expert's turns. Children in the Explain-Novice group were asked "How did you know that?" or "How could you have known that?" after each of their turns. The form of the Explain-Novice prompt was determined by their accuracy for that problem. Children in the Control group were not prompted for explanations. Prompts for explanations were always given after the correct answer had been identified and checked.

Measures

Addition skill. Each addition problem children answered was scored as correct or incorrect. Nonresponses were scored as incorrect but were very rare (< 2% of all trials). Difference scores were calculated by subtracting each child's pretest score from each of their training scores. Difference scores were selected for analysis because the selection criteria artificially reduced variance on the pretest scores, making them unsuitable to serve as a covariate.

Strategies. On the basis of videotapes of the sessions, all children's addition strategies were classified as (a) count all, (b) count-guess, or (c) retrieval/other. Counting was considered any trial in which the child overtly counted his or her fingers, counted out loud, or counted quietly while moving his or her lips. To be considered a count-all strategy, children had to deliberately attempt to represent both addends. Any counting strategy not classified as count all was count-guess. A typical count-guess strategy consisted of simply counting on fingers 1 to 10 and saying 10, with no apparent attempt to represent either addend. All noncount strategies were classified as retrieval/other.

The strategy classification scheme used in this study was developed for a prior study of children learning addition skills in collaboration with an adult (Calin-Jageman, Ratner, & Foley, 2001). Other studies have utilized a finer grain of strategy classification (e.g., Geary et al., 1991); these studies have typically benefited, however, from overtly asking children how they solved each problem. In this study, only children in the Explain-Novice group were asked to explain their thinking. Thus, a somewhat coarse grain of coding was found to be the most reliable across all experimental conditions.

Children's addition strategies were classified separately by the first author and by a second coder who was unaware of the study's hypotheses. Initial interrater reliability was very high, with a Cronbach's alpha of 0.88. Disagreements were reconciled by discussion on a trial-by-trial basis.

Encoding. Children's explanations were analyzed for recapitulation of any of three different aspects of the expert's demonstrated strategy: (a) counting in general, (b) counting two separate addends, or (c) integrating the two counts into a total sum. Children were given 1 point for each of these three elements of the expert's strategy to create a score from 0 to 3.

The first author and a second coder unaware of the this study's hypotheses classified all explanations. Initial Cronbach's alpha was 0.93. Disagreements were reconciled by discussion on a trial-by-trial basis.

Explanations. Children's explanations were classified as providing *what* information (e.g., "I counted on my fingers"), *what* + *why* information (e.g., "I counted on my fingers because it helps me"), no relevant information (e.g., "I have this game at home"), or no response. Again, both the first author and a second coder classified all explanations. Initial Cronbach's alpha was 0.88. Disagreements were reconciled by discussion on a trial-by-trial basis.

Analysis

For most dependent variables, two-way analyses of variance (ANOVAs) were conducted to analyze the effects of condition and session. Significant main effects of condition were followed up with planned comparisons contrasting each experimental group with the control group. Significant main effects of session were followed up with Bonferonni-adjusted pairwise comparisons.

RESULTS

Preliminary Analyses

Of the 39 children pretested, 28 met the criteria to complete the experiment (15 boys and 13 girls) and 11 exceeded the criteria (8 girls and 3 boys). No relation was observed between gender and qualification status (met criteria, exceeded criteria), $^2(1, N = 39) = 2.19, p < .14$. Furthermore, gender was not related to pretest performance. Boys scored an average of 23% correct ($SEM = 4.7\%$) and girls an average of 26% correct ($SEM = 4.5\%$), a nonsignificant difference, $t(37) < 1$. In all subsequent analyses, no differences between boys and girls were observed. Therefore, gender is not reported as a factor in the analyses following.

Following the pretest, children were randomly assigned to one of three experimental conditions. Pretest means for the Control Group, Explain-Novice Group and Explain-Expert Group were 15%, 17%, and 17% correct, respectively ($SEM = 5\%$ for each group). A one-way ANOVA showed that there were no significant differences in pretest performance among the three experimental groups, $F(2, 24) < 1$.

Addition Accuracy

Over the course of the experiment, children improved their ability to solve simple addition problems. Mean accuracy was 16% correct during the pretest (*SEM* = 2.8%), 20.3% during the first experimental session (*SEM* = 3.7%), and then 32%, 34%, and 40% during the second, third, and fourth experimental sessions, respectively (*SEMs* = 6%, 6%, and 7%). A one-way ANOVA across all five sessions (pretest through fourth experimental session) confirmed that accuracy improved over the course of the experiment, $F(4, 104) = 5.86$, $p < .01$.

To explore whether improvements in accuracy scores were related to experimental condition, difference scores were analyzed in a 3 (condition: Control, Explain-Novice, Explain-Expert) × 4 (session: Training 1 through 4) ANOVA (see Figure 1). The main effect of session was significant, $F(3, 72) = 3.71$, $p < .05$, indicating that improvements in accuracy occurred across the sessions. The main effect of condition was also significant, $F(2, 24) = 3.85$, $p < .05$. The interaction term was not significant, $F < 1$. Planned comparisons indicated that the Explain-Expert group improved their scores more than the Control Group ($p < .05$) and more than the Explain-Novice group ($p < .05$). The Explain-Novice group, however, was not significantly different from the Control group ($p < .81$).

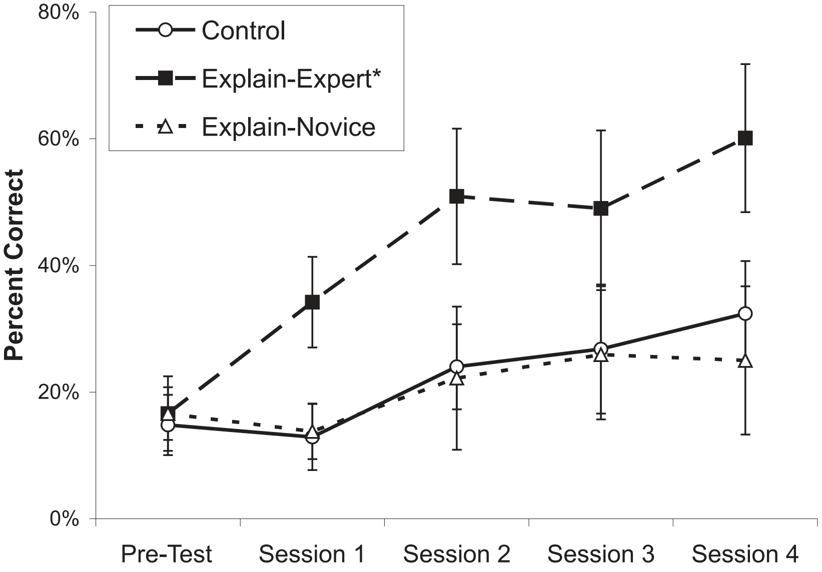

FIGURE 1 Addition skill by test session and experimental condition. Percentage correct (*M* ± *SEM*) by experimental session for the Control group ($n = 9$, solid line, circle markers), Explain-Expert group ($n = 9$, dashed line, square markers), and Explain-Novice group ($n = 9$, small dashed line, triangle markers).

Count-All Strategy

Strategy acquisition. Children were selected for not knowing the count-all strategy. Over the course of the experiment, however, many children acquired this addition strategy. By the end of the first experimental session, 25% of the children ($n = 7$) had utilized the count-all strategy at least once. This percentage grew to 44% ($n = 12$) by the end of the second session, 55% ($n = 15$) by the end of the third session, and 59% ($n = 16$) by the end of the experiment.

Experimental condition did not affect the proportion of children who acquired the count-all strategy. Of the 16 children who used the strategy at least once, six were from the Control Group, seven from the Explain-Expert group, and three from the Explain-Novice group. The association between group and acquisition (did/did-not-acquire) was noticeable, but not significant, $^{2}(2, N = 16) = 3.98, p < .13$.

Strategy use. Once children acquired the count-all strategy, they continued to use it. Of the 15 children who acquired the count-all strategy before the last session, 12 used the count-all strategy at least once in every subsequent session. The frequency with which children used the count-all strategy increased from 17% of the trials in the first experimental session (SEM = 5.7%) to 30% in the second session (SEM = 7%), and then 34% and 36% in the third and fourth sessions, respectively (SEMs = 8%, 7%). This increase within the sessions was statistically significant, as shown by a one-way ANOVA across the four sessions, $F(3, 81) = 3.63, p < .05$.

To analyze the impact of experimental condition on children's selection of the count-all strategy, count-all frequencies were analyzed with a 4 (session) × 3 (condition) repeated-measures ANOVA. Because children were selected for not knowing the count-all strategy, no adjustment for pretest scores was required. Table 1 presents group means and standard errors. The main effects of session, $F(3, 72) = 2.97, p < .05$, and condition, $F(2, 24) = 4.29, p < .05$, were significant. The interaction was not significant, $F(6, 72) = 1.07, p < .38$. Planned comparisons indicated

TABLE 1
Count-All Frequencies and Standard Errors by Session and Condition

	Training							
	1		*2*		*3*		*4*	
	%	*SE*	%	*SE*	%	*SE*	%	*SE*
Control	14	.11	21	.13	25	.15	20	.12
Explain-Other[a]	30	.10	53	.16	62	.15	65	.13
Explain-Self	8	.08	19	.11	11	.13	17	.13

[a] Overall significantly different from Control.

that children in the Explain-Expert group used the count-all strategy more frequently than children in the Control group ($p < .05$). There was no difference between children in the Explain-Novice and Control groups ($p < .87$). Following up the main effect of session, pairwise comparisons indicated that children used the count-all strategy less during the first experimental session than in any of the other sessions ($p < .05$ for each comparison); no other comparisons were significant.

Count-all and learning. Children who acquired the count-all strategy accounted for most of the learning observed in the experiment. Specifically, the 16 children who acquired the strategy improved their scores by an average of 42% by the fourth session (*SEM* = 6%); the 11 children who did not acquire the strategy improved only by 5% (*SEM* = 7%). The 16 children who eventually acquired the count-all strategy scored 16% correct on the pretest (*SEM* = 3%); the 11 children who did not acquire the strategy scored 15.9% (*SEM* = 5%), a nonsignificant difference, $t(25) < 1$. There were also no differences in initial encoding, $M = 0.56$, 0.27; $SEM = 0.15$, 0.19 for acquiring or not acquiring the strategy, respectively; $t(25) = 1.16$, $p < .25$.

Count-Guess: A Bridge to Count-All

As explained earlier, the count-guess strategy was considered counting in which no attempt was made to represent or combine addends. Typically, this involved simply counting 1 through 10 and stating 10 as an answer (48% of all count-guess trials). Sometimes, children would continue counting beyond 10 for some time and give the highest number they could think of as an answer (26% of all count-guess trials). Another type of count-guess response involved counting some or all fingers and then giving an answer apparently unrelated to the actual count. We first observed the count-guess strategy during a prior experiment that also examined the development of children's addition skills while collaborating with an expert adult (Calin-Jageman et al., 2001). The strategy suggests children's unsuccessful attempts to mimic the expert's count-all strategy. Children were about 20% accurate with the count-guess strategy—a surprisingly high percentage. The relative success of the count-guess strategy was probably due to the fact that two of the problems in the problem set summed to 10.

Though the count-guess strategy is not highly effective, it may have served as an important bridge to children's acquisition of the count-all strategy (see Figure 2). Of the 16 children who acquired the count-all strategy, 14 first used a count-guess strategy at least once. Moreover, count-guess frequency decreased after acquisition of the count-all strategy. This can be seen by examining four-trial windows before and after the trial of acquisition for the 16 children who acquired the count-all strategy. Children used the count-guess strategy in 51% of the four trials prior to acquisition (*SEM* = 9%); 14% of the four trials following acquisition

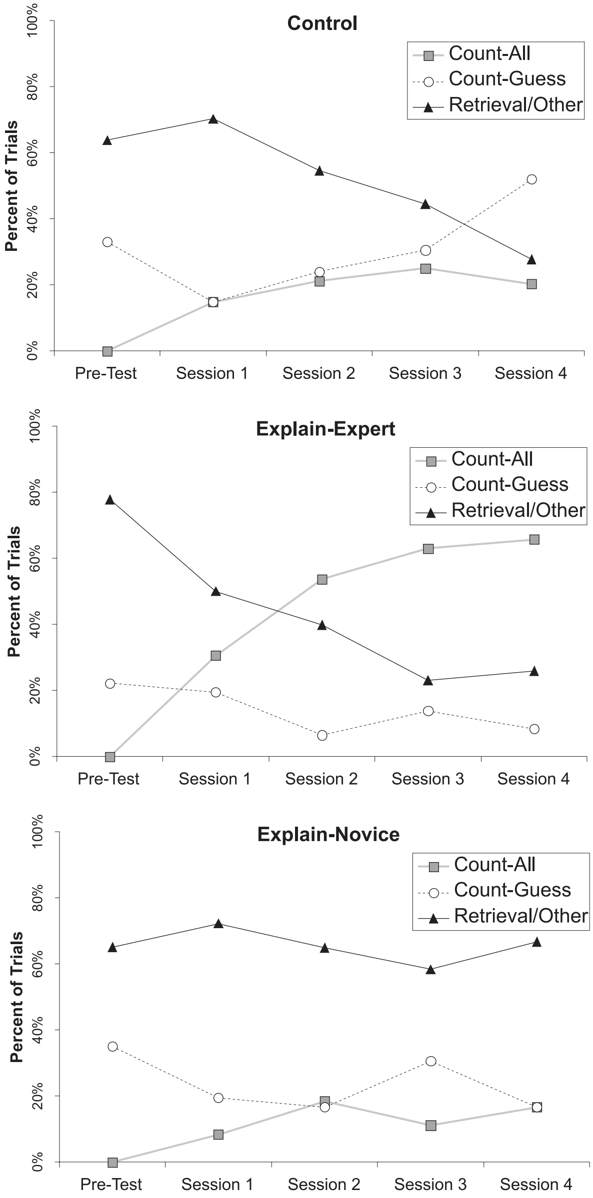

FIGURE 2 Strategy use by test session and experimental condition. Percentage of trials on which children used the count-all strategy (*solid line, squares*), count-guess strategy (*dashed line, circles*), and retrieval/other strategy (*gray line, triangles*) for each test session for the Control group ($n = 9$, top), Explain-Expert group ($n = 9$, middle), and Explain-Novice group ($n = 9$, bottom).

(*SEM* = 6%), and 15% of the fifth through eighth trials following acquisition (*SEM* = 7%). A repeated-measures ANOVA confirmed that use of the count-guess strategy decreased following count-all acquisition, $F(2, 30) = 7.68$, $p < .01$. Specifically, use of the count-guess strategy was reduced in both of the postacquisition windows relative to the preacquisition window ($p < .01$ for both comparisons). These results suggest that the count-guess strategy served as a stepping stone for children to acquire the count-all strategy. Once the count-all strategy was acquired, the count-guess strategy was largely (though not entirely) replaced.

Under the assumption that the count-guess strategy represents children's attempts to successfully master the expert's count-all strategy, the speed with which children progressed from count-guess to count-all was measured by tallying the trials between the first appearance of each strategy. Of the 14 children who used the count-guess strategy before acquiring the count-all strategy, an average of 6.07 trials separated the appearances of each strategy (*SEM* = 1.3). Thus, on average, children required about a session and a half of observing and solving addition problems to learn to complete the expert's strategy successfully. Interestingly, children in the Explain-Expert group progressed from the count-guess strategy to the count-all strategy more rapidly than children in the Control group. Children in the Explain-Expert group transitioned in an average of 3.2 trials (*SEM* = 0.8); children in the Control group in 9.6 trials (*SEM* = 2.3), a significant difference, $t(10) = 2.44$, $p < .05$. Comparisons with the Explain-Novice group were not possible, as only three children in this group acquired the count-all strategy.

Noncounting Strategies

Children used a counting strategy on 47% of all addition trials. The remaining 53% were classified as simply retrieval/other. This category could potentially combine (a) retrieval from memory, (b) guessing, and (c) covert counting. It is likely, however, that the retrieval/other classification primarily represents guessing. Children were accurate on only 17% of retrieval/other trials. Furthermore, a repeated-measures ANOVA showed that strategy efficiency did not improve over the course of the experiment ($F < 1$). Finally, children's responses on trials scored as retrieval/other did not match the problem set well. Although correct responses ranged from 6 to 10, 15% of retrieval/other responses were less than 6 and 12% were greater than 10. An additional 15% of responses were equal to the first or second addend of the problem, a common guessing strategy. Thus, it seems likely that at least 42% of strategies classified retrieval/other were actually some form of guessing.

Consistent with this interpretation, children decreased their reliance on the retrieval/other strategy as the experiment progressed. This strategy accounted for 69% of children's solutions during the pretest, 64% during the first experimental session, and then 54%, 41%, and 40% during the second, third, and fourth sessions, respectively (*SEMs* = 8% for each session). A one-way ANOVA indicated that this change was significant, $F(4, 104) = 4.66$, $p < .01$. Specifically, selection of the retrieval/other strategy was significantly reduced in the third and fourth sessions relative to baseline. Experimental condition was not related to reductions in the retrieval/other strategy, $F < 1$.

Encoding Scores

Over the course of the experiment, children improved their encoding of the expert's behavior. During the pretest, when children had observed the expert demonstrate the strategy only once, children scored an average of 0.44 (*SEM* = 0.12) out of 3 on the encoding task. Means subsequently improved to 1.07, 1.07, 1.22, and 1.37 during Experimental Sessions 1 through 4, respectively (*SEMs* = 0.14, 0.12, 0.16, 0.17). Thus, children advanced from irregularly mentioning counting (which would score a 1) to mentioning counting and the addends that were counted. A repeated-measures ANOVA confirmed that encoding scores had changed over the course of the experiment, $F(4, 104) = 9.58$, $p < .01$. Pair-wise comparisons confirmed that encoding scores in every training session were improved over the pretest ($p < .01$ for each comparison).

To analyze the impact of experimental condition on changes in encoding a 4 (session) × 3 (condition) analysis of covariance (ANCOVA) was performed with pretest encoding scores as the covariate (Table 2). The covariate was a significant predictor of later encoding scores, $F(1, 23) = 5.57$, $p < .05$, that met requirements for homogeneity of regression—Pretest × Session $F < 1$; Pretest × Condition, $F(3, 23) = 1.68$, $p = .20$. In the analysis, only the main effect of condition was significant, $F(2, 23) = 5.43$, $p < .05$. Planned comparisons indicated that the Explain-Expert group improved their encoding scores more than children in the Control ($p < .05$) and Explain-Novice groups ($p < .05$). This group difference emerged rapidly, with Explain-Expert children demonstrating increased encoding within

TABLE 2
Encoding Scores and Standard Errors by Session and Condition

			Session							
	Pretest		*1*		*2*		*3*		*4*	
	Score	*SE*	*Score*	*SE*	*Score*	*SE*	*Score*	*SE*	*Score*	*SE*
Control	0.67	.22	0.89	.25	0.67	.21	1.11	.28	1.56	.31
Explain-Expert[a]	0.33	.22	1.56	.25	1.67	.21	1.89	.28	1.67	.30
Explain-Novice	0.33	.22	0.78	.25	0.89	.21	0.67	.28	0.89	.30

[a] Overall significantly different from Control.

the first experimental session. There was no difference between the Control and Explain-Novice groups ($p < .75$).

As predicted, Explain-Expert children encoded more and learned more than children in the Control group; Explain-Novice children showed no differences with the Control group. These results suggest that the benefits of self-explaining are strongly related to the influence of self-explanation on encoding processes. This interpretation depends on the assumption that strong encoding facilitates later performance. Consistent with this interpretation, overall change in encoding scores was positively correlated with overall change in addition accuracy, $r = 0.39$, $n = 27$, $p < .05$. Furthermore, children who acquired the count-all strategy also showed the biggest changes in encoding. This was demonstrated in a 4 (session) × 2 (acquisition status) ANCOVA with pretest encoding scores serving as a covariate. Only the main effect of acquisition status was significant, $F(1, 24) = 4.30$, $p < .05$, demonstrating that children who acquired the strategy showed bigger gains in the encoding measure than children who did not.

Explanations

Children were willing to explain both their own and the expert's behavior on the treasure chest game. Only 3% of responses to explanation prompts were "I don't know" and only 1% were null responses. Despite this willingness to explain, however, children's explanations were not always interpretable or relevant (e.g., "My grandmother helps me with numbers" or "It was easy, Silly!" or "Because I can draw 9"). In fact, nearly 45% of children's explanations were classified as "relational"—as eliciting information from the child that was related to, but not actually relevant to the task at hand. Relational explanations were not analyzed further.

Of the explanations that were relevant to the task at hand, many were fragmentary and elliptical. For example, "I … I thought 1, 2, 3. the whole treasure!" or "because you put 2 in and I put 4 in … we put some treasure in." Only a small minority of explanations were well-formed and complete, for example, "first I thought only 7 so I counted and then I thought it would be 8 and I was right" or possibly "because my brain … memory!").

The fragmentary nature of kindergartner's explanations made fine-grained content analysis impossible. Instead, the broad classification scheme described earlier was adopted, in which explanations were classified as providing *what* information, *what* + *why* information, or no relevant information. For example, in the Explain-Novice group the response "cause I just. I just guessed it in my head" was classified as a *what* explanation whereas the response "counted my fingers 'cause its easier" was classified as a *what* + *why* explanation. Similarly, in the Explain-Expert condition "you counted" was classified as a *what* explanation but "fingers cause you want to get it right" was classified as a *what* + *why* explanation.

Children in the Explain-Novice group were asked “How did you know that?” or “How could you have known that?” depending on their previous answer accuracy. To ensure that the counterfactual prompt was not detrimental to children in this group, explanation content was compared for each type of prompt. Explain-Novice children generated *what* explanations for 32% of the regular prompts and 28% of the counterfactual prompts. This 4% difference was not significant, $t(7) < 1$. Children gave *what* + *why* explanations following 2% of the regular prompts and 3% of the counterfactual prompts. Again, this difference was not significant, $t(7) < 1$.

To examine the effect of experimental condition on explanation content, the percentage of *what* explanations was analyzed in a 4 (session) × 2 (condition: Explain-Novice, Explain-Expert) ANOVA. The main effect of condition was not significant, $F(1, 16) = 2.70, p < .12$. Neither the main effect of session nor the interaction were significant.

What + *why* explanations were analyzed separately from *what* explanations because the measures are not independent (every *what* + *why* explanation is also a *what* explanation). A 4 (session) × 2 (condition: Explain-Novice, Explain-Expert) ANOVA showed no significant main effects or interaction. However, the main effect of condition approached significance, $F(1, 16) = 4.09, p < .060$.

Children who produced more *what* + *why* explanations also learned more, as shown by a positive correlation with overall change in accuracy, $r = 0.48, n = 18, p < .05$. There was also a positive correlation between generating *what* + *why* explanations and changes in encoding, $r = 0.64, n = 18, p < .05$.

DISCUSSION

The results of this experiment demonstrate a strong relation between encoding and the self-explanation effect. Children who were prompted to explain an expert’s strategy (Explain-Expert) encoded the strategy better and learned more than children in the Control group. Children prompted to explain their own strategy (Explain-Novice) showed neither advantage. Thus, the power of self-explaining seems to be related to better encoding of the learning strategy. This interpretation is consistent with the large body of evidence showing that strong encoding can facilitate learning (e.g., Siegler & Chen, 1998), including the correlation observed in this study between encoding and learning. It is also consistent with the two accounts of the self-explanation effect put forward by Chi and associates (Chi, 2000; VanLehn et al., 1992), both of which predict that generating explanations enhances encoding.

This experiment also demonstrates a clear role for the content of children’s explanations. Across explanation conditions, producing *what* + *why* explanations was correlated with gains in encoding and addition skill. This is consistent with studies of both older and younger students (Chi et al., 1994; Siegler, 1995), which also show that generating *why*-type explanations can benefit learning. The efficacy of the Explain-Expert condition may be partly related to explanation content, as children in this group were somewhat more likely to produce *what* + *why* explanations than children in the Explain-Novice group.

Crowley and Siegler (1999) have previously examined how producing explanations facilitates learning. In their study, children (kindergarten–second grade) observed an expert demonstrating of a sophisticated tic-tac-toe strategy. All children were prompted to explain the expert's strategy; some children also listened to explanations of the strategy. Encoding was measured by having children predict the expert's moves and reproduce the expert's strategy. Across ages and conditions, posttest performance was best predicted by only two factors: encoding and the frequency of correctly explaining the expert's strategy. In our experiment, we reproduced this relation between encoding and learning. By including a Control group we were also able to demonstrate that this relation underlies the effectiveness of generating explanations.

Learning With an Expert Partner

Our results suggest strong links between self-explaining, encoding, and learning. What additional processes mediate these links? In particular, what processes are involved in generating explanations and how do they impact student's memories? Part of the answer may be contained in the social context that self-explaining generates between the learner and the instructor.

Across conditions, this experiment revealed a strong influence of social interaction on strategy acquisition and selection. Although children were selected for not knowing the count-all strategy, they immediately began using counting strategies after observing the expert solve problems with the count-all strategy. Specifically, almost 60% utilized a counting strategy (count-all or count-guess) by the end of just one round with the adult. By the end of the experiment, 81% had used a counting strategy at least once. This rapid shift in strategy selection indicates that children were developing and selecting strategies not only from their own performance but also through observation of the adult expert. It is important to note, however, that this shift was not a stage-like progression to the count-all strategy. Children who acquired the count-all strategy continued to make frequent use of the count-guess and noncounting strategies. This pattern of multiple strategy usage and quantitative change in strategy selection fits well with the overlapping-waves theory of development (Siegler, 1996). This approach was developed primarily from analysis of children learning on their own, but it also seems to capture the path of change in the dyadic learning task we developed for this experiment.

Although children rapidly began counting their fingers after observing the expert use the count-all strategy, the process of acquiring this strategy was complex and involved more than just rote imitation. Most children began by counting their fingers and then guessing an answer (the count-guess strategy). Although this strategy was relatively ineffective, it seemed to provide an important bridge to acquiring the count-all strategy—87% of the children who acquired the count-all strategy first used the count-guess strategy, and reliance on the count-guess strategy immediately declined on the first use of the count-all strategy. This pattern of acquisition shows that children were readily able to acquire the external aspects of the expert's behavior, but required time, effort, and continued interaction to understand and utilize the meaning of this behavior. This mix of imitation, deduction, and social interaction has been termed appropriation (Rogoff, 1990) and internalization (Vygotsky, 1978). It seems to have been the dominant form of learning in this experimental task, as the 16 children who appropriated the expert's strategy accounted for all improvement in addition skills.

Explaining, Appropriating, and Encoding

If the social interaction between child and expert encouraged appropriation of the expert's behavior, the Explain-Expert condition may have enhanced the quality of this interaction, leading to better encoding and better learning. Relative to the Control Group, children in the Explain-Expert showed two advantages in strategy acquisition: (a) They acquired the strategy more rapidly and (b) they used it more frequently after acquisition. The increased rate of acquisition suggests that generating explanations can enhance children's ability to understand and internalize the meaning of another's behavior. This makes sense, as the very nature of the explanation prompt ("How did I [the expert] know that?") is an invitation for the child to reflect on the rationale and significance of someone else's behavior and the reasoning behind it. As Chi (2000) argued, this could lead children to detect gaps in their own understanding and to infer additional information about the expert's strategy. One byproduct of this process would be enhanced encoding of the expert's strategy. Thus, appropriation may serve as the critical mediator between explaining and encoding.

After acquiring the count-all strategy, children in the Explain-Expert condition were more likely to use this strategy than children in the other experimental conditions. This suggests that generating explanations helped children exert better control over their later behavior. Specifically, children in the Explain-Expert condition were better able to resist old patterns of problem solving such as count-guess. This effect illustrates Vygotsky's (1978) claim that children turn social speech inward to organize and guide their behavior. It seems that the opportunity to speak about the expert's behavior may have given children an external pattern of behavior that could be turned inward to guide their own thought. Acquisition of this external pattern may also have guided children's ability to recall the expert's behavior, thus enhancing their performance on the encoding tests.

If explaining another's reasoning draws children more deeply into the collaborative process of appropriating an expert's strategy, this may have a measurable effect on the way children remember their interactions with the expert. Specifically, collaborative learning encourages a coordination of multiple perspectives (Tomasello, Kruger, & Ratner, 1993), a process that can blur the line between the cognitive activity of self and other (Rogoff, 1990). Explaining another's reasoning, then, may improve learning but impair children's ability to distinguish their own work from the work of the expert, even though memory of the actions within the interaction are accurate (Ratner, Foley, & Gimpert, 2002).

Educational Implications

This study continues to extend the body of literature showing that young students can benefit from generating explanations (Crowley & Siegler, 1999; Peters et al., 1999; Pine & Messer, 2000; Siegler, 1995). From our results and those of Siegler (1995) it may appear that young students would benefit most from explaining a teacher's didactic demonstrations. However, the lack of benefits for the Explain-Novice group may have been related to their complete ignorance of the count-all strategy—it is possible that students with partial knowledge of the strategy may have benefited from explaining their own reasoning. Furthermore, Siegler (2002) has provided evidence that maximal benefits are achieved when students explain an expert's correct reasoning *and* explain the errors made by novices. This approach is consistent with successful math education practices in Japan (Stigler & Hiebert, 1999), and it may help children overcome faulty reasoning while also appropriating an expert's reasoning.

Our results suggest that instructors who prompt young students to self-explain may be frustrated because over 50% of the explanations in our experiment were elliptical and off-topic. Our results were obtained, however, using scripted prompts. In the course of a natural dialogue with students, instructors may be able to elicit more on-topic explanations (though see Chi, Siler, & Jeong, 2004). In shaping student's explanations through dialogue, instructors should encourage students to focus on why a solution works, as our results and others (Chi et al., 1994) show that students benefit most from generating *why*-type explanations.

REFERENCES

Ashcraft, M. (1992). Cognitive arithmetic: A review of data and theory. *Cognition, 44,* 75–106.

Bielaczyc, K., Pirolli, P., & Brown, A. (1995). Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge acquisition activities on problem solving. *Cognition and Instruction, 13,* 221–252.

Calin-Jageman, R., Ratner, H., & Foley, M. (2001). *The self-explanation effect and children's acquisition of addition skills: The role of perspective-taking* [Poster abstract]. Conference on Human Development, Memphis, TN.

Chi, M. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology: Educational design and cognitive science* (Vol. 5, pp. 161–238). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Chi, M., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13,* 145–182.

Chi, M., De Leeuw, N., Chiu, M., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18,* 439–477.

Chi, M., Siler, S., & Jeong, H. (2004). Can tutors monitor students' understanding accurately? *Cognition and Instruction, 22,* 363–387.

Crowley, K., & Siegler, R. (1999). Explanation and generalization in young children's strategy learning. *Child Development, 70,* 304–316.

Geary, D., Brown, B., & Samaranayake, V. (1991). Cognitive addition: A short longitudinal study of strategy choice and speed-of-processing differences in normal and mathematically disabled children. *Developmental Psychology, 47,* 175–192.

Miller, P. (2000). Individual differences in children's strategic behaviors: Utilization deficiencies. *Learning and Individual Differences, 6,* 285–307.

Neuman, Y., & Schwarz, B. (1998). Is self-explanation while solving problems helpful? The case of analogical problem-solving. *British Journal of Educational Psychology, 68,* 15–24.

Neuman, Y., & Schwarz, B. (2000). Substituting one mystery for another: The role of self-explanations in solving algebra word-problems. *Learning and Instruction, 10,* 203–220.

Newman, D. (1986). The role of mutual knowledge in the development of perspective taking. *Developmental Review, 6,* 122–145.

Peters, L., Messer, D., Smith, P., & Davey, N. (1999). An investigation into Karmiloff-Smith's RR model: The effects of structured tuition. *British Journal of Developmental Psychology, 17,* 277–292.

Pine, K., & Messer, D. (2000). The effect of explaining another's actions on children's implicit theories of balance. *Cognition and Instruction, 18,* 35–51.

Ratner, H., Foley, M., & Gimpert, N. (2002). The role of collaborative planning in children's source-monitoring errors and learning, *Journal of Experimental Child Psychology, 81,* 44–73.

Ratner, H., Wendorf, C., & Hill, L. (2004). *Facilitating children's memory performance in conversation and play.* Manuscript submitted for publication.

Renkl, A., Stark, R., Gruber, H., & Mandl, H. (1998). Learning from worked-out examples: The effects of example vocabulary and elicited self-explanations. *Contemporary Educational Psychology, 23,* 90–108.

Rogoff, B. (1990). *Apprenticeship in thinking.* New York: Oxford University Press.

Siegler, R. (1995). How does change occur: A microgenetic study of number conservation. *Cognitive Psychology, 28,* 225–273.

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

Siegler, R. (2000). The rebirth of children's learning. *Child Development, 71,* 26–35.

Siegler, R. (2002). Microgenetic studies of self-explanation. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 31–58). Cambridge: Cambridge University Press.

Siegler, R., & Chen, Z. (1998). Developmental differences in rule learning: A microgenetic analysis. *Cognitive Psychology, 36,* 273–310.

Stigler, J., & Hiebert, J. (1999). *The teaching gap*. New York: Free Press.

Tomasello, M., Kruger, A., & Ratner, H. (1993). Cultural learning. *Behavioral and Brain Sciences, 16,* 495–552.

VanLehn, K., Jones, R., & Chi, M. (1992). A model of the self-explanation effect. *Journal of the Learning Sciences, 2,* 1–59.

Vygotsky, L. (1978). *Mind in society.* Cambridge, MA: Harvard University Press.