Training and Extension of a Memory Strategy: Evidence for Utilization Deficiencies in the Acquisition of an Organizational Strategy in High- and Low-IQ Children

David F. Bjorklund
Florida Atlantic University

Wolfgang Schneider
University of Würzburg

William S. Cassel and Elizabeth Ashley
Florida Atlantic University

BJORKLUND, DAVID F.; SCHNEIDER, WOLFGANG; CASSEL, WILLIAM S.; and ASHLEY, ELIZABETH. Training and Extension of a Memory Strategy: Evidence for Utilization Deficiencies in the Acquisition of an Organizational Strategy in High- and Low-IQ Children. CHILD DEVELOPMENT, 1994, 65, 951-965. 145 9- and 10-year-old children were classified into high- and low-IQ groups and given 4 different sort/recall lists (baseline, training, near [immediate] extension, far [1-week] extension) to assess training and extension of an organizational memory strategy. All children received categorized items of moderate typicality for Phases 1, 3, and 4. For Phase 2, children were assigned to either a training or control group, with half of the children in each group receiving category typical items and the others category atypical items. Levels of recall, sorting, and clustering were greater in Phase 2 for high-IQ children, for the typical lists, and for trained children. Both the high- and low-IQ children trained with typical items continued to show high levels of recall on the near extension phase. No group of subjects maintained high levels of recall after 1 week, although levels of sorting and/or clustering on the extension trials remained high for all groups of subjects except the low-IQ control children. This latter pattern (elevated sorting/clustering with low levels of recall) is an indication of a utilization deficiency, a phase in strategy development when children use a strategy but gain little or no benefit in performance. The results provide evidence for IQ, training, and material effects in the demonstration of a utilization deficiency.

There are many factors that have been found to influence individual differences in children’s memory performance, including knowledge base (e.g., Bjorklund & Berntson, 1986; Schneider & Bjorklund, 1992; Schneider, Körkel, & Weinert, 1989), metacognition (Borkowski & Peck, 1986; Schneider, Körkel, & Weinert, 1987), and intelligence (e.g., Borkowski & Peck, 1986; Scruggs & Mastropieri, 1988). Previous research examining the relation of knowledge base and intelligence (usually IQ) to memory performance has demonstrated that, in many contexts, individual differences in knowledge base can compensate for generally low academic/intellectual ability. For example, in research by Schneider et al. (1989), third-grade soccer experts had higher levels of recall from a soccer story than did fifth- and seventh-grade soccer novices, a finding similar to those of other expert/novice studies with children (e.g., Chi, 1978;)

This research was supported by a grant from the Spencer Foundation. We would like to thank the staff and students of J. C. Mitchell Elementary and A. D. Henderson University Elementary Schools for their cooperation in conducting this study. We also thank Susan Edge and Ann Santoro for assistance in collecting and analyzing data and Jane F. Gaulney, Thomas R. Coyle, Allison E. Gribbin, Louis Perry, Lenore Read, and three anonymous reviewers for comments on earlier drafts of this manuscript. Correspondence regarding this paper should be sent either to David F. Bjorklund, Department of Psychology, Florida Atlantic University, Boca Raton, FL 33431 (E-mail: BITNET: bjorkldf@fauvax; INTERNET: bjorkldf@acc.fau.edu), or to Wolfgang Schneider, Institut für Psychologie, University of Würzburg, Wittelbachplatz 1, D-8700 Würzburg, Germany (E-Mail address: wschneider@vax.rz.uni-wuerzburg.dbp.de).
Gaultney, Bjorklund, & Schneider, 1992; Opwis, Gold, Gruber, & Schneider, 1990; Recht & Leslie, 1988). Of greater significance was the fact that high- and low-IQ experts did equally well remembering the soccer story. In other words, low-IQ children showed comparable levels of recall to that of high-IQ children when tested on materials from their domain of expertise (see also Recht & Leslie, 1988; Walker, 1987).

Most research examining the relation of knowledge base and intelligence, however, has concentrated on tasks that involve little in the way of deliberate, conscious strategic functioning (e.g., text processing). Less work has been done examining the role of knowledge base on the acquisition and extension of explicitly deliberate memory strategies in high- and low-IQ children.

In fact, an examination of the developmental literature on the relation between intelligence (as reflected by IQ) and strategic memory performance fails to present a clear picture. For example, in studies examining memory-strategy training in gifted and nongifted children, gifted children usually are found to be more strategic and to demonstrate greater transfer of the strategy than nongifted children (e.g., Borkowski & Peck, 1986; Scruggs & Mastropieri, 1988; Scruggs, Mastropieri, Jorgensen, & Monson, 1986), but not always (e.g., Harnishfeger & Bjorklund, 1990; Muir, Masterson, Wiener, Lyon, & White, 1989). Within the normal range, more strategic individuals have been found regardless of IQ (e.g., Borkowski & Peck, 1986), whereas others have not (e.g., Schneider et al., 1992), or have found significant relations only under certain experimental conditions (e.g., Black & Rollins, 1982).

Although there are many differences among the studies finding and not finding significant relations between IQ and strategic memory performance, our interpretation of the literature is that high-IQ children usually have an advantage when relations among the to-be-remembered items are less obvious (i.e., categorically atypical; Harnishfeger & Bjorklund, 1990) and when transferring the strategy to new information or over substantial delays (e.g., Borkowski & Peck, 1986). Based on previous research, acquisition and extension of a memory strategy should be most likely, for both high- and low-IQ children, when highly familiar or category typical items serve as stimuli (e.g., Bjorklund, 1988; Bjorklund & Buchanan, 1989; Hasselhorn, 1992; Rabinowitz, 1984, 1988). However, it is likely that low-IQ children will be most apt to acquire and extend a strategy only for the more typical items. Such a tendency, if it were found, would argue for the use of familiar and well-integrated concepts as instructional materials, particularly for low-achieving students.

The focus of the present study was individual differences in children's abilities to extend a trained memory strategy as a function of the typicality of the to-be-remembered materials. More specifically, we classified children into high- and low-IQ groups and gave some of them training in the use of an organizational strategy. We also varied the typicality of the items on which children received training.

Nine- and 10-year-old children, classified as either high or low IQ, were administered four sort/recall trials, using different sets of items on each trial. Nine- and 10-year-old children were selected for this study because children of this age rarely use an organizational strategy spontaneously with sets of category atypical items, but can readily do so with relatively simple training (e.g., Bjorklund & Buchanan, 1989). There is also much variability in spontaneous strategy use among children of these ages, which we believe would make contrasts between high- and low-IQ children fruitful.

All subjects received sets of moderately typical items in Phases 1 (baseline), 3 (near extension), and 4 (far extension—1 week after training). In the second phase, half of the children in each IQ group received training in the use of an organizational strategy, whereas control subjects merely sorted and recalled items as in the other phases. During this second phase, half of all children in each of the four training × IQ groups received sets of category atypical items and half received sets of category typical items. We predicted that the high-IQ children would benefit from training with both the typical and atypical items and demonstrate near and possibly far extension of the trained strategy. We also predicted that high-IQ subjects in the control condition would benefit from receiving the typical items in Phase 2 and extend a discovered strategy to new sets of items. In contrast, low-IQ children were hypothesized to demonstrate extension effects only under highly facilitative circumstances. Thus, although the low-IQ children were
predicted to experience enhanced recall due to (a) training and (b) exposure to the typical items in the absence of training, they were expected to demonstrate extension of the trained strategy only under the most supportive condition (i.e., training with typical items).

A secondary purpose of this study was to examine patterns of changes in recall, sorting, and clustering over phases to get a better picture of the dynamics of strategy acquisition in children. Production deficiencies have long been part of the cognitive developmental literature (e.g., Flavell, 1970), reflecting children's tendencies not to use a strategy spontaneously, even though they can benefit from the strategy when given proper instruction. More recently, Miller (1990) has suggested an additional phase in strategy acquisition—that of utilization deficiency. According to Miller (1990), a utilization deficiency is a transitional phase in strategy development, between no strategy use and the efficient (and beneficial) implementation of a strategy. During this time, strategies are used ineffectively. Children who show a utilization deficiency will spontaneously generate a strategy but may not gain any performance benefit from it, demonstrate less benefit than that shown by older children, or may even experience a decline in performance. Several experiments have demonstrated utilization deficiencies in preschool children (e.g., Baker-Ward, Ornstein, & Holden, 1984; DeMarie-Dreblow & Miller, 1988; Miller & Harris, 1988; Miller, Haynes, DeMarie-Dreblow, & Woody-Ramsey, 1986), and one recent study has reported a utilization deficiency in third-grade children for a spontaneously generated organizational memory strategy in a free-recall task (Bjorklund, Coyle, & Gaultney, 1992; see also Masur, McIntyre, & Flavell, 1973, who reported no benefit from the use of a selective study strategy for 9-year-olds). In the present study, we examined changes in children's strategy use, as reflected by patterns of sorting and clustering, and related them to changes in recall, looking for evidence of a utilization deficiency. Unlike previous studies examining utilization deficiencies, the present experiment includes multiple measures of strategy use (sorting and clustering) and assesses a variety of factors that may influence the likelihood of a child's displaying a utilization deficiency (e.g., intelligence, typicality of materials). It thus represents an extension of the earlier literature, tempting to move beyond demonstration to an explanation of the factors that affect children's strategy use and their effectiveness.

Method

Design

Nine- and 10-year-old children were classified as high or low IQ and assigned to one of four training x typicality conditions. All children received four sort/recall trials with categorically related items, using different sets of items and categories in each phase (trial). All children received sets of moderately typical items in Phase 1 (baseline), Phase 3 (near extension), and Phase 4 (far extension). In Phase 2, half of the children received sets of category typical items and the remaining half received sets of category atypical items. Within each typicality condition, half of the children received training in the use of an organizational strategy during Phase 2 (trained subjects), with the remaining (control) subjects receiving standard sort/recall instructions, as in the other three phases. Phases 1, 2, and 3 were conducted in a single session, with Phase 4 being run 1 week later. This resulted in a 2 (IQ group) x 2 (training condition) x 2 (typicality) x 4 (phase) design, with repeated measures on the phase factor.

Subjects

Subjects in this experiment were 62 third graders (32 boys and 30 girls, mean age = 9.17 years, SD = .57 years) from five different classrooms and 81 fourth graders (37 boys and 44 girls, mean age = 10.18 years, SD = .38 years) from six different classrooms. Children were selected from two public schools in south Florida and came from a wide range of socioeconomic backgrounds.

Children were classified into high- and low-IQ groups, based on median splits of the scores on the combined Vocabulary and Information subtests of the WISC-R, calculated separately at the third and fourth grades. The mean estimated verbal IQ of the low-IQ children was 87.2 (SD = 12.5); the mean estimated verbal IQ of the high-IQ children was 111.8 (SD = 7.5).

Materials, Tasks, and Procedures

All children were seen individually and were given one free-recall trial in each of four phases. On each trial, children were presented with a series of items from natural language categories prepared as line drawings on 3 x 3-inch cards, which had previously been judged as either moderately
TABLE 1
ITEMS USED IN THE EXPERIMENT

Typical Sets (Used in Phase 2 Only)

Set A: FRUIT: apple, peach, pear, banana, grapes, strawberry
CLOTHING: shirt, pants, dress, socks, coat, shorts
FURNITURE: chair, couch, table, bed, desk, lamp

Set B: VEGETABLES: cucumber, carrot, lettuce, tomato, potato, celery
BODY PARTS: leg, arm, foot, ear, hand, neck
TRANSPORTATION: car, bus, truck, airplane, bicycle, train

Atypical Items (Used in Phase 2 Only)

Set A: FRUITS: lime, coconut, tangerine, mango, prune, apricot
CLOTHING: belt, vest, hat, shoes, scarf, gloves
FURNITURE: fan, clock, rug, nightstand, stereo, bench

Set B: VEGETABLES: mushroom, onion, squash, turnip, eggplant, pumpkin
BODY PARTS: hair, jaw, ankle, elbow, shoulder, heel
TRANSPORTATION: tricycle, tractor, skates, spaceship, wagon, helicopter

Moderately Typical Sets (Used in Phases 1, 3, and 4 Only)

Set A: TOOLS: nails, axe, screwdriver, lawn mower, saw, drill
MUSICAL INSTRUMENTS: flute, violin, harmonica, banjo, trumpet, drum
BIRDS: eagle, sparrow, cardinal, owl, chicken, robin

Set B: WEAPONS: sword, knife, cannon, arrow, club, rifle
SPORTS: basketball, soccer, volleyball, hockey, tennis, golf
MAMMALS: lion, wolf, zebra, pig, giraffe, rabbit

Set C: DOMICILES: tent, trailer, igloo, cabin, teepee, castle
OCCUPATIONS: teacher, cashier, cook, artist, soldier, nurse
TOYS: puzzle, jump rope, marbles, balloon, crayons, swing

Practice Items (Used in Training)

Set A: VEGETABLES: peas, beans, sweet potato
BODY PARTS: fingers, wrist, nose
TRANSPORTATION: jeep, taxi, skateboard

Set B: FRUITS: cherries, pear, pineapple
CLOTHING: blouse, t-shirt, tie
FURNITURE: couch, bookcase, stool

NOTE.—The typical and atypical sets were used in Phase 2 only. Sets A, B, and C of the moderately typical items were used with equal frequency in Phases 1, 3, and 4. Practice set A was used in training for subjects who received typical or atypical set A; practice set B was used in training for subjects who received typical or atypical set B.

typical, highly typical, or atypical, based, collectively, on norms generated by adults (Rosch, 1975; Uyeda & Mandler, 1980) and children (Bjorklund, Thompson, & Ornstein, 1983). Each list contained three categories with six items per category. The items used in phases 1, 3, and 4 were all of moderate typicality, whereas the lists used in Phase 2 consisted of either category typical or atypical items. All lists were balanced for word frequency (Thorndike & Lorge, 1944). The order in which the moderately typical lists were presented was counterbalanced over Phases 1, 3, and 4. Two sets of category typical and two sets of category atypical lists were used during Phase 2, counterbalanced across subjects. The items used in this experiment are presented in Table 1.

The moderately typical lists consisted of items from a broad range of typicality. Thus, unlike the typical or atypical lists, which were homogeneous with respect to the category typicality of their items, the moderately typical lists were constructed to be more heterogeneous with respect to category typicality and similar to categorical lists traditionally used in developmental memory experiments.
In a first session, children were seen individually by one of three female experimenters and given three sort/recall trials with different sets of items (Phases 1, 2, and 3). Children were told that they would be shown some pictures that they were to study and to "do whatever you want with them in order to remember them." Children were told that, later, the experimenter would ask them to remember the pictures in any order they liked. Children were also shown examples of the Matching Familiar Figures Test (MFFT) and told that they would be asked to solve some problems on this test later.

During Phase 1 (baseline), children were shown pictures of moderately typical items. The experimenter showed each card to the children, at a rate of about one card every 2 sec, and named the item, with children repeating the name after the experimenter. Cards were placed in front of the children in a $3 \times 6$ matrix in one of two predetermined random orders, with the stipulation that no two items from the same category be presented consecutively. Following presentation, children were told again that they could do anything they wanted with the cards to help them remember and that they would later be asked to recall them in any order they liked. After 2 min, the cards were covered with an opaque cloth. As a buffer-clearing task, children were directed to solve problems on the MFFT for approximately 30 sec. Children were then asked to recall as many items as they could in any order they wished. If a child was silent for 10 sec, the experimenter asked if there were any additional items that he or she could remember. When either another 10-sec interval occurred with no items recalled or the child stated that he or she could not remember more items, the trial ended.

For subjects in the control condition, Phase 2 immediately followed the completion of Phase 1 with the same procedure, the only exception being that subjects received sets of category typical or atypical items. Subjects in the trained condition were given a detailed instructional exercise using one of two practice sets of nine new items from three new categories (see Table 1). The experimenter explained to children that "a good trick to use when remembering things is to sort them into categories of items that go or belong together." They were also given the rationale that it is "usually easier to remember one large category than several small details" and that the larger category would remind them of the items in that category. This was accompanied by a demonstration of how to sort the cards and by labeling the categories used. The children were given a chance to practice the new method with the nine practice items and received corrective feedback on their performance. They were encouraged to remember what they had been shown and were told "don't forget to use your new skill the next time you want to remember things." Next, a new set of 18 items (either typical or atypical) was presented as in Phase 1. The children were given an opportunity to sort and study the cards for 2 min. As before, the cards were then covered and a buffer-clearing task administered; the remainder of the phase was identical to Phase 1.

All subjects were treated identically in Phase 3 (near extension), receiving sets of moderately typical items and the same instructions as those given in Phase 1. Subjects in the trained condition were not reminded to use their new "skill" in Phase 3, but merely told, as were the control subjects, to "study the cards and do whatever you want with them in order to remember them later." At the conclusion of this phase, both the Vocabulary and Information subtests of the WISC-R were administered to subjects.

Approximately 7 days after the initial session, children were tested again to investigate delayed extension of training (far extension). Phase 4 instruction were the same as those for Phase 1 with a new set of 18 moderately typical items.

Results

All effects are reported at $p < .05$, with significant effects being evaluated by Newman-Keuls tests, unless otherwise stated.

---

2 At the conclusion of the first session, children were asked several metamemory questions, specifically inquiring about what strategies they had used during the task. Group differences in professing a strategy were found, with no differences in the distribution of subjects professing a strategy among the high-IQ trained (89%), low-IQ trained (82%), and high-IQ control (94%) subjects, although each of these groups was more likely to profess using a strategy than the low-IQ control subjects (55%), $\chi^2(1) \approx 6.06$. Because of ceiling levels of strategy profession for most groups of children, the results of the metamemory interview did not affect the interpretation of the other results, and they will not be discussed further.
1a - High-IQ, Trained, Typical

1b - High-IQ, Trained, Atypical

1c - Low-IQ, Trained, Typical

1d - Low-IQ, Trained, Atypical
Fig. 1.—Mean recall, sorting, and clustering over phases for each condition × IQ level × typicality group. Recall (maximum = 18) indicated on left-hand Y-axis; sorting (ARC) and clustering (ratio of repetitions) indicated on right-hand Y-axis. (Note: Phase 1 = baseline; Phase 2 = training; Phase 3 = near extension; Phase 4 = far extension. All subjects received items of moderate typicality in phases 1, 3, and 4. Subjects received typical or atypical items in Phase 2 only.)
Recall

Mean recall (maximum = 18) is presented by IQ group (high vs. low), condition (trained vs. control), typicality (typical in Phase 2 vs. atypical in Phase 2), and phase (baseline, training, near extension, far extension) in Figure 1a–1h. Analysis of these data produced significant main effects of IQ group, F(1, 135) = 16.69, MSE = 22.23 (high IQ, M = 10.81 > low IQ, M = 9.19), condition, F(1, 135) = 5.46 (trained, M = 10.5 > control, M = 9.5), and phase, F(3, 405) = 22.36, MSE = 4.18. Also significant were the condition × phase, F(3, 405) = 3.71, and the typicality × phase, F(3, 405) = 12.18, interactions.

Examination of the significant condition × phase interaction demonstrated that recall for control subjects was significantly enhanced only in Phase 2 (M = 10.23), with recall at Phases 1 (M = 9.03), Phase 3 (M = 9.52), and Phase 4 (M = 9.23) being comparable. For subjects in the trained condition, recall at Phase 2 (M = 12.11) was significantly greater than recall at each of the other three phases, although recall at Phase 3 (M = 10.67) was significantly greater than that at Phase 4 (M = 9.64) and Phase 1 (M = 9.58). Recall differences between the trained and control groups were significant at Phases 2 and 3 only, ts(540) ≥ 2.33. This interaction reflects the fact that training resulted in improved levels of recall, which were maintained on a near extension trial.3

Subsequent analysis of the significant typicality × phase interaction showed no significant differences in recall across the four phases for subjects receiving the atypical items during Phase 2 (mean recall = 9.98, 10.31, 9.96, and 9.93 for Phases 1 through 4, respectively). In contrast, subjects receiving typical items during Phase 2 showed significantly greater recall at Phase 2 (M = 11.92) than at all other phases, with recall at Phase 3 (M = 10.20) being significantly greater than recall at Phase 4 (M = 8.99) and Phase 1 (M = 8.86). This pattern of results indicates that receiving the typical items during Phase 2 resulted in significant enhancement of recall and was maintained on a near extension trial. Moreover, enhancement was found when subjects received typical items during Phase 2 independent of training condition.

Sorting

The degree to which subjects organized their sorts according to adult-defined categories was assessed using the Adjusted Ratio of Clustering (ARC) measure (Roenker, Thompson, & Brown, 1971), and can be seen in Figure 1a–1h. An ARC score of 1.0 reflects perfect sorting, with zero representing chance, and negative values reflecting sorting less than that expected by chance. Subjects who did not sort items at all during the sort phase were given a sorting score of 0.

The analysis of the sorting data produced significant main effects of IQ group, F(1, 135) = 9.23, MSE = .04 (high IQ, M = .57 > low IQ, M = .36), condition, F(1, 135) = 62.87 (trained, M = .69 > control, M = .28), and phase, F(3, 405) = 80.9, MSE = .08. Also significant was the condition × phase interaction, F(3, 405) = 36.36.

Examination of the significant condition × phase interaction showed that sorting increased slightly over trials for subjects in the control condition, with sorting at Phases 3 (M = .31) and 4 (M = .37) being significantly greater than sorting at Phase 1 (M = .18), and with sorting at Phase 2 (M = .25) being undifferentiated from sorting at Phases 1 and 3 but significantly less than sorting at Phase 4. For subjects in the trained condition, sorting increased substantially from Phase 1 (M = .17) to Phase 2 (M = .66), with levels of sorting remaining high at Phases 3 (M = .85) and 4 (M = .87). Differences in sorting between the trained and control conditions were nonsignificant at Phase 1, but were significant at each of the remaining three phases, ts(540) ≥ 6.53.

For many subjects, sorting according to the predefined categories (see Table 1) was perfect once they discovered the technique of grouping items. Table 2 presents the percentage of subjects in each condition × typicality × IQ group who sorted perfectly (i.e., ARC = 1.0) for each phase. Although few subjects in any group sorted perfectly in Phase 1, the percentage of trained subjects

3 Although the condition × phase effect did not interact significantly with the typicality and IQ factors, an examination of Figure 1d clearly shows that the low-IQ subjects trained on atypical items did not maintain their advantage on the near extension trial, but actually demonstrated a slight decline in recall relative to Phase 1 (i.e., 9.87 vs. 9.71). Moreover, recall for the high-IQ subjects trained on the atypical items (see Fig. 1b) was only slightly greater than that observed in Phase 1 (11.63 vs. 10.63), indicating that near extension effects were most prominent for subjects trained with the typical items in Phase 2.
who sorted perfectly in Phase 2 exceeded 80% in each group, and was 39% (compared to 22% at baseline) for high-IQ control subjects receiving the typical items. The percentage of subjects continuing to sort perfectly remained high in Phases 3 and 4 for all groups of trained subjects and increased, relative to baseline, for the high-IQ control subjects receiving both the typical and atypical items. These patterns illustrate the fact that many subjects did acquire an organizational strategy and employed it perfectly on extension trials. What is of further interest is that perfect sorting on the extension trials was often associated with relatively low levels of recall, comparable to that found at baseline.

Clustering

Clustering in recall was evaluated using the ratio of repetition score, defined as $r/n - 1$, where $r$ refers to the number of intracategory repetitions and $n$ refers to the number of items recalled (Bousfield, 1953). The ratio of repetition score was chosen because it does not vary systematically with level of recall (Murphy, 1979). Mean clustering scores are presented in Figure 1a–1h.

Analysis of the clustering data yielded significant main effects of IQ group, $F(1, 135) = 17.41, MS_e = .08$ (high IQ, $M = .58 >$ low IQ, $M = .49$), condition, $F(1, 135) = 22.27$ (trained, $M = .60 >$ control, $M = .48$), typicality, $F(1, 135) = 3.88$ (typical, $M = .56 >$ atypical, $M = .52$), and phase, $F(3, 405) = 60.10, MS_e = .03$. Also significant were the following interactions: condition $\times$ phase, $F(3, 405) = 6.69$, typicality $\times$ phase, $F(3, 405) = 7.79$, IQ $\times$ phase, $F(3, 405) = 3.44$, and IQ $\times$ condition $\times$ phase, $F(3, 405) = 3.29$.

Subsequent analysis of the typicality $\times$ phase interaction revealed that clustering was greater for subjects given the typical items only at Phase 2 (typical, $M = .69 >$ atypical, $M = .55$, $t(540) = 4.01$). Examination of the IQ group $\times$ condition $\times$ phase interaction showed that clustering increased from Phase 1 to Phase 2 for all groups of subjects, although the increase was considerably larger for the trained than for the control children. Levels of clustering remained high, and significantly greater than at Phase 1, for both the high- and low-IQ groups who received training (see Fig. 1a and 1b; 1c and 1d), and for the high-IQ control subjects (see Fig. 1e and 1f). The pattern was different for the low-IQ control subjects, whose clustering was comparable at Phases 1, 3, and 4 (see Fig. 1g and 1h).

Differences in Patterns of Recall for Comparably Strategic Children

Research has generally reported that brighter children are more apt to use strategies than less bright children. The greater tendencies of higher IQ children to use strategies, in turn, have been proposed to be responsible for their enhanced levels of task performance. The question remains, however, whether children of different levels of intelligence will perform comparably given equivalent levels of strategy use. That is, will lower IQ children show as much benefit from using a strategy as higher IQ children?

We assessed this question by contrasting the recall of children who showed perfect sorting (i.e., ARC = 1.0) in Phases 2 (training), 3 (near extension), and 4 (far extension). Separate IQ group $\times$ typicality analyses of variance were performed at each phase, including as subjects only those chil-
Mean recall × IQ level for Phases 2, 3, and 4 are shown in Table 3. As can be seen, mean recall was greater, overall, for high-IQ relative to low-IQ children at each phase. Results of the analyses demonstrated that this difference was significant at Phase 3, $F(1, 74) = 5.53$, $MSe = 8.67$, and approached significance at Phase 2, $F(1, 70) = 3.02$, $p < .10$, $MSe = 7.36$. The effect of IQ was not significant at Phase 4, $F(1, 81) = 1.03$, $MSe = 8.41$, $p > .10$. The effect of typicality was significant only at Phase 2, $F(1, 70) = 11.04$, where subjects received different sets of materials. There were no other significant effects in any of the analyses. Thus, despite comparable and perfect sorting, levels of recall were generally higher for the brighter children.

Classifying Subjects as Utilizationally Deficient

The above analyses demonstrate different patterns of performance across the four phases of this experiment for recall and the two measures of strategy use, sorting and clustering. For many groups of subjects, levels of recall on the extension phases were low and comparable to those observed on the baseline phase, whereas levels of sorting and/or clustering remained high. Such a pattern is indicative of a utilization deficiency, with subjects using a strategy (as reflected either by sorting or clustering) but showing no benefits in recall performance. Although these group data are informative, group data can mask patterns of individual performance, and it would be useful to classify subjects individually as to whether they display a utilization deficiency.

We modified a procedure used by Bjorklund, Coyle, and Gaultney (1992) to classify subjects individually as utilizationally deficient, separately on Phases 3 and 4, for both sorting and clustering. We classified a subject as utilizationally deficient (i.e., using a strategy without corresponding improvement in levels of recall) if his or her levels of sorting or clustering were significantly greater on Phase 3 or Phase 4 than they were on Phase 1 (baseline), without a two-word or more improvement in recall. We also examined the distribution of subjects who were classified as utilizationally deficient on both the near and far extension phases, separately for sorting and clustering (see Bjorklund, Coyle, & Gaultney [1992] for a detailed rationale for and description of the procedure).

To determine if levels of sorting or clustering were significantly greater on Phases 3 or 4 relative to Phase 1, we computed the minimal difference between two means required for significance, based on Newman-Keuls tests (three steps, $p = .01$), using the error term from the appropriate analysis of variance. These values were .29 for sorting (ARC scores) and .18 for clustering (ratio of repetitions scores).

An examination of sorting revealed that few control subjects were classified as utilizationally deficient, in part because of the relatively low frequency of control subjects showing perfect sorting (see Table 2). For the trained subjects, 37% were classified as utilizationally deficient on Phase 3, 49% on Phase 4, and 34% on both Phases 3 and 4. The distribution of subjects classified as utilizationally deficient on Phases 3 and 4 for the sorting measure did not vary as a function of which items they received in Phase 2 (i.e., typical vs. atypical) or IQ group, all $\chi^2(1) \leq 1.99$.

An examination of clustering revealed that slightly more trained subjects were classified as utilizationally deficient than control.

---

### Table 3

**Mean Recall for Subjects with Perfect Sorting (ARC = 1.0) by IQ Group Separately for Phases 2, 3, and 4**

<table>
<thead>
<tr>
<th>IQ Group</th>
<th>Phase 2</th>
<th>Phase 3</th>
<th>Phase 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>High IQ</td>
<td>12.39</td>
<td>12.29</td>
<td>10.92</td>
</tr>
<tr>
<td>Low IQ</td>
<td>11.73</td>
<td>11.36</td>
<td>10.17</td>
</tr>
</tbody>
</table>

---

4 We performed similar analyses for the second strategy measure, clustering. Perfect clustering was much less frequent than perfect sorting, however, and so we arbitrarily chose a ratio of repetition score of .70 as our criterion. In contrast to the sorting data, there were fewer subjects meeting this criterion. Although fewer effects reached significance in this analysis, the pattern of data was similar to that reported for the sorting scores, and because of space limitation, these data will not be provided here.
for all groups of subjects. Recall being comparable to that of Phase 1 extension of recall, however, with Phase 4 atypical) items. There was no evidence of far extension task. These results support earlier findings indicating that strategies are more easily acquired, trained, and extended by lower-ability children for implementing a strategy is so great as to leave too few resources to devote to other aspects of the task until that strategy can be executed with greater efficiency (see Bjorklund & Harnishfeger, 1987). Thus, increased task perfor-

Although the sorting and clustering measures produced slightly different patterns of results, most trained subjects who were classified as utilitarian deficient for one measure were similarly classified as utilitarian deficient for the other measure. The overall percentages of trained subjects classified as utilitarian deficient for both sorting and clustering were 29% for Phase 3, 30% for Phase 4, and 21% for both Phases 3 and 4, values that are only slightly lower than those obtained for the separate sorting and clustering measures. Stated another way, given that a subject was classified as utilitarian deficient for either sorting or clustering, 73% were similarly classified as utilitarian deficient on the second measure for Phase 3, 59% for Phase 4, and 64% for both Phases 3 and 4.

Discussion

As expected, levels of memory performance, sorting, and clustering were overall greater for high-IQ than for low-IQ children, higher on lists of typical than atypical materials, enhanced as a result of training, and, under some conditions, generalized on a near extension task. These results support earlier findings indicating that strategies are more easily acquired, trained, and extended on typical than atypical sets of items (e.g., Bjorklund, 1988; Bjorklund & Buchanan, 1989; Hasselhorn, 1992; Rabinowitz, 1984, 1988). They further demonstrate that low-IQ children can substantially enhance their performance as a result of training with typical items and extend that training to a new set of less typical items. High-IQ children do the same, but will also acquire and extend an organizational strategy with more exposure to sets of typical (and to a lesser extent, atypical) items. There was no evidence of far extension of recall, however, with Phase 4 recall being comparable to that of Phase 1 for all groups of subjects.

One of the most interesting findings of this study was that indices of strategy use—sorting and clustering—did not follow patterns of recall, but varied as a function of IQ, training, typicality of the items, and phase of the experiment. For all groups of trained subjects, both near and far extension were found for both the sorting and clustering measures. The high-IQ control children, however, demonstrated significant near and far extension effects only for clustering, not for sorting. The low-IQ control subjects were the only group that did not show an extension effect for either sorting or clustering. Although Bjorklund, Coyle, and Gaultney (1992) had previously demonstrated enhanced levels of clustering over trials in the absence of enhanced levels of recall for school-age children, in the present study, strategy use was indicated by sorting as well as clustering, yielding a measure of strategic functioning that is independent of recall (unlike clustering). In fact, as can be seen from Table 2, a majority of the trained subjects and about half of the high-IQ control subjects displayed perfect sorting on the far extension trial, despite the fact that no group of subjects showed an extension effect for recall in Phase 4.

The nonlinear relation between strategy use and recall was further illustrated by the presence of IQ-related differences in memory performance among equally strategic groups of children. When examining only the data of children who were perfectly strategic in sorting (i.e., sorted items according to adult-defined categories), higher IQ children had higher levels of recall than lower IQ children. Thus, given comparable and perfect strategy use, a recall advantage still persisted for the brighter children. These findings are consistent with other research that has reported higher levels of memory performance in higher relative to lower IQ children in the absence of differences in strategy use (e.g., Harnishfeger & Bjorklund, 1990; Muir et al., 1989). Apparently, factors other than strategy use, such as differences in knowledge base or speed of processing (e.g., Bjorklund, Coyle, & Gaultney, 1992; Kail, 1991), continue to influence task performance. In fact, it is possible that the additional mental effort required by lower-ability children for implementing a strategy is so great as to leave too few resources to devote to other aspects of the task until that strategy can be executed with greater efficiency (see Bjorklund & Harnishfeger, 1987). Thus, increased task perfor-
Performance should not be viewed as a necessary consequence of using a strategy, even a strategy that has been shown to be effective when used by more intellectually advanced children.

Although relations between recall, sorting, and clustering did vary across the different groups and phases, the predominant pattern was for elevated levels of sorting and/or clustering in the absence of elevated levels of recall. Such a pattern reflects a utilization deficiency, which has been demonstrated by Miller and her colleagues for young children for simple strategies (e.g., Miller, 1990), and more recently by Bjorklund, Coyle, and Gaultney (1992) for older children for an organizational memory strategy. The current results extend these findings, taking the concept of utilization deficiency beyond the demonstration phase, by identifying conditions under which a utilization deficiency occurs and showing that utilization deficiencies occur not only for spontaneously produced strategies (as in the earlier studies and for subjects in the control condition here), but also for quasi-spontaneous strategies (i.e., spontaneous production of a strategy after training).

Factors Affecting Utilization Deficiency

Training.—First of all, our data indicate that utilization deficiencies are more apt to be found following training. Both the high- and low-IQ trained subjects showed patterns of utilization deficiency on Phases 3 and 4 (i.e., stable recall with elevated levels of sorting and clustering), indicating that children from a wide range of intellectual abilities will extend a strategy, often without recall benefits, following training. In contrast, control subjects were less likely to display a utilization deficiency, for the obvious reason that, without the benefit of instruction, they were less likely to implement a strategy.

IQ.—Second, our results indicate that high-IQ children are more likely to be classified as utilizationally deficient than low-IQ children. Because high-IQ children are more likely to be spontaneously strategic than low-IQ children (subjects cannot be classified as utilizationally deficient unless they produce a strategy), it is under the control conditions of this experiment that IQ effects in utilization deficiency should most likely be found, and were in our data. These results suggest that, for a given strategy, brighter children may not only be the first to use a strategy, but the first to experience a utilization deficiency.

Typicality of items.—The likelihood of being classified as utilizationally deficient also varied with the typicality of the items. Children who showed the greatest recall improvements in Phase 2 were those receiving the typical items. In fact, subjects trained with the atypical items demonstrated about a one-word gain in recall (mean advantage = 0.97 words) relative to baseline, far less than the almost four-word gain (mean advantage = 3.92 words) experienced by subjects trained with the typical words. Thus, strategy use for subjects trained with the atypical items, reflected by enhanced levels of sorting and clustering, was only marginally associated with improvements in memory performance, reflecting a mini-utilization deficiency during the training phase (see Bjorklund & Harnishfeger, 1987).

One possible interpretation of the differential effects of strategy use on recall is that the strategy was most easily implemented with the typical items used during Phase 2, and that using the same strategy with the moderately typical items during Phases 3 and 4 was more effortful, leaving fewer resources to devote to other aspects of the task (see Bjorklund & Harnishfeger, 1987; Miller, Seier, Probert, & Aloise, 1991; Miller, Woody-Ramsey, & Aloise, 1991). Such an interpretation acknowledges the difference that subjects' familiarity with the list materials has on the effectiveness of strategy use and is consistent with past theorizing (e.g., Bjorklund, Muir-Broadus, & Schneider, 1990; Folds, Footo, Guttenberg, & Ornstein, 1990; Ornstein, Baker-Ward, & Naus, 1988). Alternatively, children may have been required to devote more mental effort to accessing the strategy in Phases 3 and 4 relative to training, when a strategy was provided for them, again resulting in fewer resources available to task execution.

The discrepancy between the efficacy of strategy use during Phases 2 and 3 within an individual also illustrates that a child can use a strategy effectively in one context but use that same strategy ineffectively in another, highly similar context. In the present situation, children failed to benefit from using a strategy that had helped them just 5 minutes earlier! Thus, not only are there intraindividual differences in strategy use (e.g., McGilly & Siegler, 1989), but there are intraindividual differences in the effectiveness of the same strategy over similar conditions, indicating further the need for caution in making any inclusive statements about
How "strategic" any given child is at any given time.

Why Do Children Use an Ineffective Strategy?

One obvious question about utilizationally deficient children is, Why should they use a strategy that does not have a payoff in terms of task performance? Although the purpose of cognitive strategies is to enhance performance, they are effortful cognitive operations, and their use may not always facilitate task performance (see Bjorklund et al., 1990; Miller, Seier, Probert, & Aloise, 1991; Miller, Woody-Ramsey, & Aloise, 1991). Yet, children continued to use the effortful strategy over repeated trials without immediate benefit. We speculate that children's relatively poor metacognition may play a role in utilization deficiency. For example, Bjorklund, Coyle, and Gaultney (1992) suggested that many children in their study may not have realized that the strategy they were using was not helping them. They may have believed that doing "something" or "thinking" about how to solve a problem was generally more advantageous than doing "nothing" or not "thinking" about a problem. Given this perspective, once a strategy is discovered they may have applied it and practiced it, being unaware that it was not resulting in elevated levels of performance. Thus, a child with poor metacognition may persist in using a strategy, ultimately leading to more efficiency and improved task performance; in contrast, a more metacognitively enlightened child may recognize the current futility of using the strategy, stop using it, and actually be less likely to develop efficient strategic functioning. In this case, children's immature cognition may be seen as adaptive, leading to eventual (though not immediate) benefits (Bjorklund & Green, 1992).

Another reason why children may continue to use a currently ineffective strategy may be related to its novelty. For example, Siegler and Jenkins (1989), examining the development of simple arithmetic strategies in preschoolers, have proposed that young children may try a new strategy (such as min) despite their having repeated success with a simpler strategy (such as sum), in part because of the novelty of the former strategy. Taking a cue from Piaget (1970), Siegler and Jenkins proposed that people are often interested in exercising newly acquired cognitive abilities. Something similar may be happening with utilization deficient children, who may discover a new strategy and use it because of the novelty. With repeated use, the novelty wears off; however, by that time, the strategy may be yielding improved levels of performance, which will cause it to be chosen over other, less efficient strategies (see Siegler & Jenkins, 1989).

Unfortunately, we have no data from this study to assess either of these hypotheses, although recent research with preschool children provides some tentative support for the cognitive immaturity idea. In a study examining preschool children's abilities to evaluate their imitative skills, 3- and 4-year-olds who were "out of touch" with their imitative abilities (i.e., who overpredicted how well they could imitate an adult model) scored higher on a verbal IQ subtest than 3- and 4-year-old children who were more accurate in their predictions. The relation was reversed for 5-year-olds (Bjorklund, Gaultney, & Green, 1993). In other words, for young subjects, brighter children were less metacognitively aware than less bright children. Although these findings relate only indirectly to the relation between metacognition and utilization deficiencies, they are provocative and suggest that searching for such a relation is worthwhile.

References


Bjorklund, D. F., Gaultney, J. F., & Green, B. L. (1993). "I watch, therefore I can do": The development of meta-imitation during the preschool years and the advantage of optimism about one's imitative skills. In M. L. Howe & R. Pasnak (Eds.), Emerging themes in cog-


This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.