Research Alive: All Our Children Learning

Jack Snowman
Southern Illinois University, Carbondale

Follow this and additional works at: https://scholarworks.bgsu.edu/mwer

How does access to this work benefit you? Let us know!

Recommended Citation
Available at: https://scholarworks.bgsu.edu/mwer/vol7/iss4/6

This Featured Article is brought to you for free and open access by the Journals at ScholarWorks@BGSU. It has been accepted for inclusion in Mid-Western Educational Researcher by an authorized editor of ScholarWorks@BGSU.
Research Alive: All Our Children Learning

Jack Snowman, Southern Illinois University, Carbondale

The title of this column was, as I'm sure many of you recognize, taken verbatim from the title of a book written by Ben Bloom and published in 1981. I chose this title for two reasons. First, each of the six recently published articles that are summarized below deals with one aspect or another of children's classroom learning. The children who participated in these studies ranged from about five years of age (kindergarten) to roughly thirteen years of age (eighth-grade). Second, and more importantly, each of these studies is consistent with the basic reason why Bloom wrote All Our Children Learning. On page 1 of the Introduction, Bloom states that as a result of research on what he called alterable variables, "student learning can be improved greatly, and it is possible to describe the favorable learning conditions which can enable virtually all students to learn to a high standard. In their own way, each of the following studies describes those learning conditions.

How to Grow a Vocabulary: Read to Children Early and Often

In volume 86, number 1 (1994) of the Journal of Educational Psychology, Claudia Robbins and Linnex Ehri make the argument that because vocabulary size is strongly correlated with children's overall school achievement, it is important to understand how young children achieve their vocabulary growth. While it might be reasonable to expect most of this growth to come from structured classroom lessons, previous research suggests that direct instruction has a negligible effect. In one study cited by these authors, 27 fourth graders learned an average of only 85 targeted words in 19 weeks. Other research suggests that listening to stories has a much stronger impact on vocabulary growth than does direct instruction. But because of shortcomings of one sort or another in these earlier studies, a definitive conclusion about story reading as a direct cause of vocabulary growth in young children could not be made.

In the Robbins and Ehri study, thirty-three non-reading kindergarten children were given the Peabody Picture Vocabulary Test-Revised (PPVT-R) and classified as having either low, medium, or high levels of vocabulary knowledge. Children in each group were randomly assigned to hear either A Crocodile's Tale or The Little Boy Who Cried Wolf. Eleven target words thought to be unfamiliar to the children were substituted for familiar words or phrases in each story. Although these words were not defined, their meanings could be obtained from cues in the surrounding context. Each story was heard twice in the space of anywhere from two to four days. After listening to the story, children took a 22-item multiple-choice vocabulary test. Eleven of the items came from the story they heard and 11 came from the story they did not hear. For each item, there were four picture choices and a "don't know" option.

Children's performance on the unfamiliar words heard in the story reliably exceeded performance on the words not heard, indicating that listening to stories was an effective means of expanding word knowledge. But the size of the difference varied as a function of PPVT-R score: children who possessed larger vocabularies to start with learned more new vocabulary words than did children with smaller vocabularies. Although the effect of hearing stories on recognition of the meanings of unfamiliar words was not large (children recognized an average of 1.24, or 11.3%, more definitions of words from the story they heard than from the story they did not hear), two things should be pointed out. First, the 11 target words were rather difficult. They included such words as decrepit, procure, query, abode, and chorale. Second, the cost of obtaining this result was almost nil. How much time and effort does it take to read a story or two to a child each day?

I Get By With a Little Help From My Friends

Recently, I asked my computer to search the ERIC database for articles dealing with cooperative learning for the years 1966 through 1993. To my great surprise it found a total of 2,712 articles. With such a large research base, one would think that all of the significant questions concerning cooperative learning have at least been asked, if not answered. Not so according to Amalya Nativ of Eastern New Mexico University. In volume 94, number 3 (1994) of The Elementary School Journal, she reports on a study that sought to determine (a) if the known relationship between helping behaviors in cooperative groups and achievement gains exists among third-, fourth-, and fifth-grade math students, (b) if gender, grade level, or ability level of the students relate to gains in achievement, (c) if all helping behaviors relate to gains in achievement, and (d) if all students exhibit taught helping behaviors.

Each cooperative learning group included students of both genders, three math ability levels (based on standardized test scores), and various ethnic backgrounds. All students were given three weeks of instruction in how to help their team members work through math problems via direct instruction, role playing, modeling, calling attention to appropriate helping behaviors, and giving points to groups whose members exhibited helping behaviors. Amalya used the cooperative learning format
known as Student Teams-Achievement Divisions (STAD). The STAD format includes direct instruction of new material, the formation of diverse student teams whose purpose is to prepare its members to do well on exams, individual improvement scores, and team recognition. Additional components for this study were initial team building, instruction on helping behaviors, and rotating students through different roles (e.g., solving problems, checking, coaching). Each team was videotaped an average of twice a week for five minutes each time.

The data analysis allowed Amalya to draw the following conclusions: (a) There is a positive and statistically significant relationship between helping behaviors in cooperative groups and subsequent math achievement for 3rd-, 4th-, and 5th-grade males and females. (2) Different helping behaviors have different effects on achievement. Of the eight helping behaviors that were examined ("received no help after requesting it," "gives explanations," "receives explanations," "receives other help," "asks for help and receives it," "gives answers only," "receives answers only"), "receives no help after requesting it" was the most influential, accounting for 9.4% of the variance in the math achievement measure, and was negatively related to achievement. The next most influential behaviors were "gives explanations" (3.9% of the variance), and "receives explanations" (2.6% of the variance). "Receives other help," and "gives other help," taken together explained less than 1% of the variance in the dependent measure. The last three behaviors were unrelated to math achievement. (3) Neither gender nor grade level was related to helping behavior, but ability was related. Students classified as high ability gave more explanations, other help, and more answers than did low ability students. (4) Although all students exhibited the helping behaviors they were taught, the more desired behaviors, such as "gives explanations," "receives explanations," and "asks for help and receives it," occurred more often than the less desired "receives no help after requesting it," "gives answers only," and "receives answers only." The Elusive Transfer Effect

Past research tells us that learning disabled (LD) students do not perform as well in school as non-learning disabled students because of deficits in such processes as encoding, conscious use of tactics and strategies, and transfer of learned information and skill. To some degree, these deficits can be remediated by teaching LD students techniques for enhancing meaningfulness, distinctiveness of encoding, and retrieval of desired information. For tasks that involve learning to associate one concept with another, the keyword mnemonic is a particularly effective device. As I have indicated in a previous column, and as Barbara Fulk, Margo Mastropieri, and Thomas Scruggs point out in volume 7, number 2 (1992) of Learning Disabilities Research and Practice, students quickly become proficient users of the keyword technique and significantly improve their memory performance in comparison to a variety of other instructional treatments. About the only desired outcome of keyword instruction that has largely escaped researchers is spontaneous transfer (also known as generalization). In fact, in an article titled "Mnemonic Strategies and Classroom Learning: A Twenty-Year Report Card (Elementary School Journal, volume 94, number 2), Joel Levin awards this area of mnemonics research a grade of D. In their study, Fulk et al. examined how well keyword instruction and attribution training helped LD students learn the meanings of unfamiliar terms from different content areas and then transfer that skill to new words lists one day and two weeks after the end of training.

Fifty-six sixth-, seventh-, and eighth-grade LD students were randomly assigned to either a mnemonic generalization training condition, a mnemonic generalization plus attribution training condition, or a rehearsal condition. The experiment consisted of three phases. Phase 1 involved one day of training with 12 words from different subject matter areas (e.g., social studies and science) and an immediate recall test. Phase 2 involved two days of practice with 10 words from different subject matter areas. Phase 3 was an unprompted generalization task that was administered one day and two weeks after the end of Phase 2.

During Phase 1 (Day 1), the mnemonic generalization training group was given explicit instructions in how to use the keyword mnemonic and was shown a set of 12 cards, each of which contained a target word (e.g., vituperation), a keyword (e.g., viper), and an illustration that incorporated the keyword and the definition (e.g., a viper speaking abusively to someone). The mnemonic generalization plus attribution training group received the same instructions and materials as did the mnemonic generalization training group. In addition, they saw two cartoon drawings, each of which was followed by a positive attribution message. The first cartoon depicted a student who had experienced success on a school task. This was followed by an index card which read: "Two main reasons why students usually do well on school tasks are these: (1) because they know a good way to remember important information, and (2) because they try hard as they use that method." The second cartoon depicted a student who had experienced failure and was followed by the message, "Two main reasons why students usually don’t do well on school tasks are these: (1) because they don’t know a good way to remember important information, and (2) because they don’t try hard.” Students in the rehearsal condition received the same set of learning materials for Phase 1 as did the other two groups. But instead of keyword instructions and illustrations, they received instructions in how to use a rehearsal tactic and simpler illustrations. For example, the card for "vituperation" depicted one person telling another, "Don’t speak so abusively to me," with
the words "vituperation" and "scolding" on the top. During Phase 2, students received (along with the set of 10 practice words) a printed card that read, "1) repeat the word, 2) study the picture of the answer information, and 3) say the word and definition together." At the end of Phase 1, each group was tested for recall with production items (e.g., "What does vituperation mean?") and matching items.

During Phase 2 (Days 2 and 3), each group reviewed the steps involved in implementing its particular technique, was told the purpose of the training, was reminded of the type of task for which the technique was appropriate, and was tested for recall with production items and matching items. The mnemonic generalization training subjects received a set of 10 cards, each of which contained a target word, a definition, and the prompt, "Did you use the strategy?" Although the students had to generate their own keywords and interactive illustrations during this phase, one card listed the steps they had been trained to use and another card listed the criteria for an appropriate keyword. The mnemonic generalization plus attribution training subjects also practiced the keyword mnemonic on a set of 10 words, but received strategy use and attribution prompts (e.g., "Did you try hard and use the strategy?") Students in the rehearsal group were given explicit explanations, prompting, and feedback of the rehearsal technique.

For Phase 3, the generalization testing, students were given new lists and told to study them with "the method that would best help them prepare" for the quiz that would follow.

The results of the training regimen described above can be fairly described as mixed. No differences were found among the groups for either of the Phase 1 recall scores. For the matching test administered after the first day of Phase 2 training, only the mnemonic generalization training group significantly outscored the rehearsal group. For the production test, both the mnemonic generalization training group and the mnemonic generalization plus attribution training group significantly outscored the rehearsal group. The scores of the two training groups did not differ significantly from each other. This last finding was replicated for both the matching and production tests administered after the second day of Phase 2 training. The one-day delayed transfer test produced no significant differences among the three groups. The two-week delayed transfer test produced no significant differences for the production test. On the matching test, however, the two training groups significantly outscored the rehearsal group.

Want to Learn More? It's a Question of Elaboration

Current cognitive conceptions of learning suggest that students learn more of a body of knowledge as they construct relationships among the ideas that make up that knowledge base and relate those ideas to what they already know. This phenomenon, which is generally known as generative learning, has a broad base of research support that stretches back over several decades. One technique that has proven effective in inducing generative learning is elaborative interrogation. Previous research has shown that elaborative interrogation, the act of generating inferences while asking why questions, increases associations among and retention for facts from a reading passage for learners from 4th-grade through college. But, as Timothy Seifert points out in volume 85, number 4 (1993) of the Journal of Educational Psychology, these effects were obtained from materials (sentences and paragraphs, for the most part) that were easily reduced to a list of equally important facts because they lacked the superordinate-subordinate structure so typical of academic prose. The study reported on by Seifert was designed to investigate whether the elaborative interrogation effect would be as strong when it was directed only at the main idea of a standard paragraph. Because a main idea is naturally elaborated by supporting subordinate ideas, and because these supporting ideas may activate relevant prior knowledge, additional elaboration may be largely superfluous. A second concern driving Seifert's study was that the specific characteristics of the elaboration itself may influence memory for the target information. Data from a small number of previous studies suggest that why questions that lead to a correct explanation of some fact are more likely to improve recall of that information than questions that lead to nonexplanatory or incorrect explanations. For example, in response to the question, "Why would American woodcocks need to be fast and unpredictable?", a correct explanation would mention that some animals protect themselves by their ability to escape from enemies. An incorrect explanation would sound plausible (e.g., to catch prey), but be wrong. A nonexplanatory response would be vague or incorrect (e.g., to fly over lakes and rivers).

To test these two hypotheses, sixth- and seventh-grade students read two passages, performed either an elaborative or non-elaborative activity while reading the second passage, and were tested for associative memory and memory for details. The first reading passage provided students with general knowledge of how animals adapt to their environment, and was intended to serve as a basis for the elaborative activity they would be asked to perform while reading the second passage. The second passage described the characteristics and behaviors of three animals. Students assigned to the underline-only group found alongside each paragraph the prompt, "Underline the most important idea in this paragraph." Students assigned to the underline-with-elaboration group were also directed to underline the most important idea from each paragraph. But the paragraphs in their passage contained an extra sentence that linked the underlined information to a particular principle of adaptation mentioned in the first passage. Students in the generate-elaboration group were instructed to answer the why...
question that appeared alongside each paragraph (e.g., "Why would the snowshoe hare need to be white in color during winter?"). Students in the elaborate-with-study-sheet group answered the same questions as did those in the generate-elaboration group, but were also given the written summaries they had made of the first passage to use as a reference.

Because Seifert was interested in the effect of elaboration versus non-elaboration, his analysis compared the performance of the three elaboration groups with the performance of the underline-only group. Differences among the three elaboration groups were not examined. Only the generate-elaboration group (those who answered why questions) significantly outperformed the underline-only group on the associative memory test, and this was entirely due, as expected, to this elaboration group's performance on main idea questions. There were no statistically significant differences between any of the elaboration groups and the underline-only group on memory for details. As far as how well each group was able to relate the specific information from the second passage to the more general information from the first passage, all three elaboration groups significantly outperformed the underline-only group. Moreover, the effect sizes were moderately strong (ES = .85, .65, and .74 for the generate-elaboration, elaborate-with-study-sheet, and underline-with-elaboration groups, respectively). In order to determine whether correct explanations given in response to why questions improved memory for factual information, a conditional probability analysis was conducted. This analysis did not find a statistically significant difference between the elaboration-with-study-sheet and generate-elaboration groups and the underline-only group.

These findings support the hypothesis that students who engage in interrogative elaborative activity while they read learn more main idea information than students who engage in non-elaborative activities, and that such activity is more likely to stimulate recall of relevant prior information. Additionally, such activity does not appear to interfere with recall of supporting details.

Retaining Students in Grade Gets a Failing Grade (Again)

In an earlier Research Alive column (Summer, 1992), I pointed out that the practice of requiring students to repeat a grade because of unsatisfactory achievement continues to be a popular one despite an almost unbroken stream of negative effects that have appeared in the research literature over the past 20 years. Samuel J. Meisels and Fong-Ruey Liaw provide yet another piece of damning evidence in volume 87, number 2 (1993) of the Journal of Educational Research. Meisels and Liaw undertook their analysis because much of the prior research was conducted on relatively small samples and there was little data on the issue of early retention (grades K-3) versus late retention (grades 4-8). Both of these problems were addressed by analyzing a subset of data from the large, representative National Education Longitudinal Study of 1988 (NELS: 88), a government-sponsored longitudinal study of student outcomes. The sample analyzed by Meisels and Liaw consisted of 16,623 eighth-grade students, 2,073 of whom were retained in grades K-3 and 1,128 of whom were retained in grades 4-8.

In response to the question, Who gets retained?, Meisels and Liaw found wide discrepancies on the basis of racial/ethnic background, gender, and social class. Thirty percent of all Black students, 25% of all Hispanic students, and 17% of all White students were retained at some point between kindergarten and eighth-grade. Twenty-five percent of males were retained as compared to 15% of females. Thirty-four percent of students from the lowest SES quartiles were retained as compared to 9% of students from the highest SES quartile. Although the overall retention rate was greater for Blacks and males, early retaineres were more likely to be White and female (despite the fact that early retention has a stronger negative impact on females than it does on males).

The timing of the retention did not produce consistent effects. Early retainees outsored later retaineres on some cognitive, emotional, and behavioral measures, but not on other measures.

Meisels and Liaw were also interested in knowing whether retention produces any academic benefits for students since this is an oft-cited rationale for invoking the practice. After controlling for students' gender, racial/ethnic background, SES, and maternal education, retained students had significantly lower grades and test scores by eighth-grade than did non-retained students. Also, retained students were about six times more likely than non-retained students to have learning problems and be assigned to special education classes, and were about twice as likely to have emotional or behavioral problems. But because the NELS: 88 data-set did not include scores on intellectual ability, Meisels and Liaw were unable to statistically control for that factor, and could not draw a definitive conclusion about the subsequent negative effects of retention. But analyses of other, albeit smaller and less representative, data sets that have controlled for initial ability, have reported negative effects on subsequent academic performance as a result of retention. The overall message from 20 years of research seems clear: grade retention is a practice that is inconsistent with, if not antithetical to, the philosophy of maximizing classroom learning.

Adolescents in the Classroom: Better Management Through Behavioral Engineering

As if the negative findings reported by Meisels and Liaw on the effect of retention aren't bad enough, there is at least one other undesirable consequence of the practice; students who are a year or two older than their classmates tend to become behavior problems. This
was the situation in which educators from the Charles­
ton (MD) County School District found themselves in
the mid-1980s. As standards for promotion were in­
creased because of a state-wide reform effort, the per­
centage of eighth-grade students who were at least one
year behind grade level increased from 34% in 1983 to
48% in 1988. Indices of student misconduct rose in con­
text with the retention rate. To help teachers deal with
and reduce student misbehavior, Denise Gottfredson,
Gary Gottfredson, and Lois Hybl designed a behavior
management program that drew heavily from operant
conditioning principles. (Despite being either ignored
or maligned as the “empty organism” theory, this writer
believes it offers sound and useful guidelines for help­
ing students acquire a variety of new behaviors.) The
program was evaluated in eight middle schools over a
three-year period. The results of their evaluation were
reported in volume 30, number 1 (1993) of the Ameri­
can Educational Research Journal.

The program was composed of four components:
establishment of a school discipline policy, a behavior
tracking system, teacher training in classroom organi­
zation and management, and teacher training in the ad­
ministration of positive reinforcement. The school dis­
cipline policy that each school adopted had to clearly
state rules for conduct, specify the punishing conse­
quences for infractions of the rules, and specify positive
consequences for exhibiting desired behavior. The be­
havior tracking system was a computer-based record of
each student’s positive and negative behaviors. Parents
were kept informed of the status of their child’s record
and encouraged to follow through at home on whatever
actions were taken at school. Teachers were shown how
to improve their classroom organization and manage­
ment practices by keeping high traffic areas free of con­
gestion, making sure that all students could be seen by
the teacher, making frequently used materials readily
available, and clearly stating rules for behavior. (Read­
ers who are familiar with the classroom management
literature may recognize that this component was taken
directly from the work of Emmer, Evertson, Sanford,
Clements, and Worsham.) Last, teachers were trained
in the appropriate use of positive reinforcement.

Because of differences in how the program was
communicated to teaching staff by school administra­
tors, transfers of administrative staff out of the target
school, and differences in attitude toward the program
on the part of key administrators, the researchers divided
the eight schools into high, medium, and low implemen­
tation categories, with the latter category serving as a
control group. The effect of the program was evaluated
by comparing how students and teachers responded to
several sets of rating scales administered at the begin­
ning of the 1987 school year and at the end of the 1989
school year. These included the Classroom Environment
Assessment (CEA), teachers’ ratings of students in-class
behavior, and the Effective School Battery (ESB).

Significant changes appeared in four of the five
CEA ratings and, with one exception, only in the high
implementation schools. Students (but, interestingly, not
teachers) reported an increase in perceived classroom
order (effect size = 0.62). Classroom organization, rule
clarity, and teacher support were also judged as having
increased (effect sizes of 0.33, 0.55, and 0.35, respec­
tively). Rule clarity was also felt to have increased in
the medium implementation school.

When teachers rated students’ on-task behav­
iors, only teachers in the high implementation schools
felt that students paid more attention to their academic
work and disrupted the classroom less frequently (effect
sizes = 0.09 and -0.12, respectively).

The statistically significant changes in ratings
by students of their own misbehavior and of the rewards
and punishments they received were more broad-based.
Increases in self-reported rebellious behavior and receiv­
ing rewards were reported by students in all schools, al­
though the increase in rebellious behavior was smallest
in the high implementation schools. Only students in
the high implementation schools reported receiving fewer
punishments.

Finally, only students in the high implementa­
tion schools reported an increase in the fairness of rules
(effect size = 0.52).