JOURNAL OF EDUCATION FOR STUDENTS PLACED AT RISK, 7(2), 241–263

The BIG Accommodation Model: The Direct Instruction Model for Secondary Schools

Bonita J. Grossen
University of Oregon

Research during the last 2 decades has found the BIG Accommodation Model to be successful in helping students with disabilities achieve rigorous academic standards in mainstream classrooms at the secondary level. This research project tested the effects of the BIG Accommodation Model in accelerating the learning of highly at-risk students in low-achieving secondary schools. The project began with 1 highly problematic middle school in Sacramento, CA and resulted in record gains for all ethnic groups, for English language learners, and for students performing at all levels in the areas of both language arts and mathematics. The project further established a professional development model that allowed for replication of success in other middle schools. Successful replication depends on the presence of three components: the curricular materials designed around "big ideas," electronic progress monitoring, and in-class coaching. The initial training model provides a particularly cost-efficient means for achieving high-quality implementations.

During the last 2 decades, research and development efforts in education have focused on the goal of closing the achievement gap between students with disabilities and general education students. The Direct Instruction (DI) model has emerged as one of the most successful models for accomplishing this goal. Table 1 highlights the findings of recent research studies evaluating whether students with disabilities are able to achieve rigorous standards in general education classes.

The remarkable successes documented in Table 1 were achieved primarily through unique engineering of the BIG Accommodation curricular materials. Table 2 contrasts traditional teacher-directed instruction with the six major instructional engineering principles used in "The BIG Accommodation" (Carnine, 1994), designed to accommodate diverse learning needs. These principles of instructional design are described in more detail in Kameenui and Carnine (2001).

Requests for reprints should be sent to Bonnie Grossen, 292 West 12th Street, Eugene, OR 97401. E-mail: bgrossen@oregon.uoregon.edu

TABLE 1
Research on the Effects of Instruction With Accommodations in Closing the Achievement Gap Between Special Education and General Education Students

Area	*Description*
Reasoning	1. On a variety of measures of argument construction and critique, high school students with learning disabilities scored as high as high school students in an honors English class and higher than college students enrolled in a teacher certification program (Grossen & Carnine, 1990) 2. In constructing arguments, high school students with disabilities scored significantly higher than college students enrolled in a teacher certification program and scored at the same level as a group of college students enrolled in a logic class (Collins & Carnine, 1988).
Science	3. On a test of problem solving to achieve better health, high school students with disabilities scored significantly higher than nondisabled students who had completed a traditional high school health class (Woodward, Carnine, & Gersten, 1988). 4. On a test of problem solving that required applying theoretical knowledge and predicting results based on given information, mainstreamed middle school students with disabilities scored higher than a class of general education students taught in a school-centered treatment (Grossen, Carnine, & Lee, 1996). 5. On a test of misconceptions in earth science, mainstreamed middle school students with learning disabilities showed better conceptual understanding than Harvard graduates interviewed in the film, *A Private Universe* (Schneps, 1987). 6. On a test of earth science problem solving, mainstreamed middle school students with learning disabilities scored significantly higher than nondisabled students who received traditional science instruction (Woodward & Noell, 1992).

(continued)

TABLE 1 (*Continued*)

Area	*Description*
	7. On a test of problem solving involving earth science content, most of a group of mainstreamed middle school students with learning disabilities scored higher than the mean score of the nondisabled control students (Niedelman, 1992).
	8. On an advanced chemistry test, high school students with disabilities and remedial students scored higher than a group of high-performing students in an advanced placement chemistry class on the subscale covering chemical equilibrium. The students matched the performance of the advanced placement students on subscales requiring application of concepts of chemical bonding, atomic structure, organic compounds, and energy of activation (Hofmeister, Engelmann, Carnine, 1989).
Mathematics	9. On a test of problem solving requiring the use of ratios and proportions, mainstreamed high school students with disabilities scored as well as nondisabled high school students who received traditional math instruction (Moore & Carnine, 1989).
	10. On a test requiring the application of fractions, decimals, and percents, age-grouped fifth and sixth grade low-achieving students scored significantly higher than high-achieving students learning in a constructivist treatment (Grossen & Ewing, 1996).
History	11. On a history test that required analyzing primary source documents, the scores that mainstreamed high school students with learning disabilities attained on the use of principles and facts in writing did not differ significantly from nondisabled control students (Carnine, Caros, Crawford, Hollenbeck, & Harniss, 1996).
	12. Middle-school urban children of poverty and some with limited English increased their history vocabulary proficiency at a rate five times that of suburban middle-school students (Carnine et al., 1996).

TABLE 2
The Contrast Between Instruction With Accommodations for Diverse Learners and Traditional Instruction

Principles of Accommodation for Diverse Learners	*Traditional Instruction*
Presentation of big ideas, concepts, and principles that facilitate the most efficient and broad acquisition of knowledge across a range of examples. Big ideas make it possible for students to learn the most and to learn it as efficiently as possible, because "small" ideas can often be best understood in relation to larger, "umbrella concepts."	Presentation of a barrage of unrelated facts and details. The links between concepts are obscured.
Teaching of conspicuous strategies made up of specific steps that lead to solving complex problems.	Strategies are seldom taught.
Background knowledge is pretaught.	Important prerequisite learning is often neither evaluated nor taught.
Mediated scaffolding provides personal guidance, assistance, and support that gradually fades as students become more proficient and independent.	Little direction or provision for scaffolding the progression of learning toward greater independence is provided.
Judicious review requires students to draw on and apply previously learned knowledge over time.	Review is often minimal.
Strategic integration blends new knowledge with old knowledge to build big ideas.	Spiraling of topics does not carefully integrate units.

BIG IDEAS: GETTING MORE ACHIEVEMENT FROM LESS LEARNING

Big Ideas in Remedial DI Programs

All the BIG Accommodation programs are new generation DI programs organized around "big ideas." Big ideas yield more power from less learning time and are the key to accelerated learning. The older generation DI programs for remediating skill deficits of secondary students provide simpler examples of big ideas. For example, the big idea in *Corrective Reading* (Engelmann et al., 1999) is the relation between sounds and letters in words. Through learning 57 sound–symbol relations (the big idea), students learn to read all the words in the language. The 57 sound–symbol relations have been carefully thought out to maximize generalizability. For example, the sound taught for the letter y is "yee," which works both at the beginning and end of words: yellow (yee-ellow), puppy (pupp-yee). And, it works in the middle of words: gym (g-yee-m). Students who are taught using *Corrective Reading* (Engelmann et al.) grow in reading at two or three times the normal rate, making it possible for many students to catch up in 1 year of instruction (for a review of the research, see Grossen, 1998).

Another remedial program, *Spelling Through Morphographs* (Dixon & Engelmann, 1999), teaches 600 morphemes and three rules for connecting them, enabling students to spell 12,000 words. The following are examples of four sets of morphemes, which combine to form four words. All four words have the morpheme "ion" in common:

1. In fect ion.
2. In flate ion (drop the final "e" when the next morpheme begins with a vowel).
3. In tent ion.
4. Pro fess ion.

Notice how the confusion in deciding whether to use "sion" or "tion" is eliminated with this analysis. If the speller considers the root meaning, the spelling is clear.

Big Ideas in Teaching Cognitively Complex Content

The new generation DI programs in the BIG Accommodation model teach cognitively complex content using the big idea analysis. Some examples follow.

Reasoning and writing. A big idea in the reasoning and writing (Engelmann & Grossen, 2001a, 2001b) program is a ruling-out process for constructing knowledge. This ruling-out process is the essence of the scientific method. Figure 1 illustrates one of the early tasks. Students must figure out what is in the mystery box and write a paragraph describing their thinking process. The outline diagram provides a template for their paragraphs. The icons graphically represent the type of thinking involved. The trapezoid prompts a summary statement, or topic sentence. The boxes illustrate the stepwise nature of the ruling-out process used in constructing knowledge. And finally, another upside-down trapezoid indicates a concluding sentence. To figure out the mystery object, students read the first clue, "The object is red," and then review the possibilities. Following the outline diagram, they write, "Clue A rules out the banana. That object is not red," and so on.

This thinking strategy has wide applications. Figure 2 illustrates the application of the ruling-out process to shopping. Henry needs a jacket and has several requirements. In this scenario there is a jacket that meets his requirements. In other activities, the students also encounter scenarios where no option meets all the requirements; they must weigh the alternatives and choose the best option. Students use this same ruling-out process for many other kinds of applications. For example, they use it to select the best plan for accomplishing a goal.

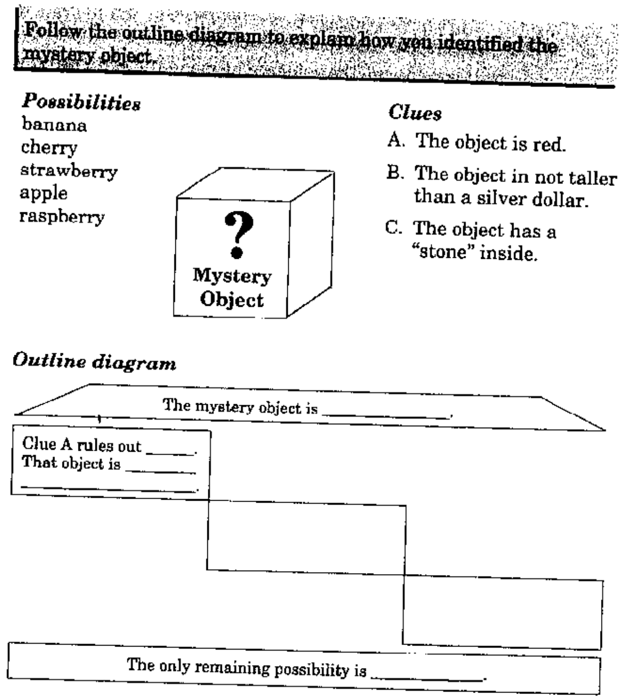


FIGURE 1 The mystery box.

This ruling-out process also represents the fundamental thinking involved in setting up and interpreting the outcomes of scientific experiments. Figure 3 illustrates a problem requiring an experiment before a conclusion can be made. Not all the possible explanations for an observation have been ruled out. The students describe a short experiment and then describe how to interpret the data, depending on how the experiment turns out. This experiment will rule out remaining possible explanations for the observation.

The outline diagrams provided in Figures 1 through 3 provide students with a clear model of the wording and thinking processes involved. Later, these prompts are faded, after students have internalized the thinking patterns and are able to work successfully without the prompts.

Earth science. The central big idea of the earth science videodisc program (Systems Impact Incorporated, 1987) is convection. Most textbooks describe convection with only one paragraph, so students never learn that convection is the basis for making predictions in earth science. For example, meteorologists predict the weather based in large part on their knowledge of convection. Woodward (1994) compared the videodisc program organized around the big idea of convection with

Part A

Henry's requirements

1. The jacket must cost less than $200.00.
2. The jacket must be washable.
3. The jacket must offer superior protection against the cold.
4. The jacket must weigh no more than 4 pounds.

Facts

Jacket	Stormbuster	Windblaster	Leader	King Kold	Wilderness
Price	$179.00	$187.99	$156.00	$206.00	$187.00
Weight	4 lb.	3 lb. 2 oz.	2 lb. 8 oz.	3 lb. 7 oz.	4 lb. 3 oz.
[illegible]	superior	superior	good	superior	superior
[illegible]	washable	dry clean only	washable	washable	washable

Outline diagram

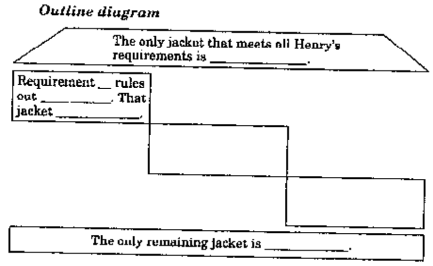
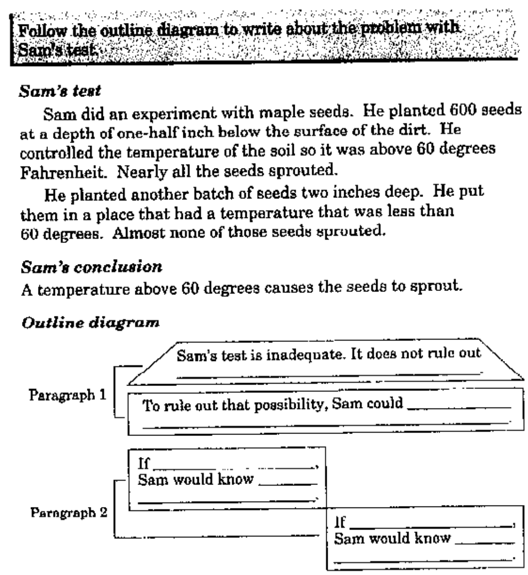


FIGURE 2 Henry's shopping problem.

a videodisc program with a standard unit organization; he found that organization around the big idea resulted in better learning, both qualitatively and quantitatively.

Understanding U.S. history. The text, *Understanding U.S. History* (Carnine, Crawford, Harniss, & Hollenbeck, 1998), uses the problem–solution–effect big idea to organize the events of history. Students learn that history is built around attempts to solve problems among groups of people. Every solution to a problem generally leads to a new problem. For example, the automobile solved a transportation problem, but created a new problem: pollution. History, therefore, can be characterized as a chain of problems. By studying the ways humans have solved problems in the past, and the effects of those solutions, students seem better able to identify solutions for the future, while attempting to avoid the mistakes of the past.

Bigger ideas integrate across content areas. The programs and practices in the BIG Accommodation provide for transfer across subject areas, thus maximizing instructional efficiency. Students learn important skills for processing, critiquing, and researching information, that they apply in all subject areas. They learn, for example, that opinions must be based on evidence, that the evidence must



Follow the outline diagram to write about the problem with Sam's test.

Sam's test

Sam did an experiment with maple seeds. He planted 600 seeds at a depth of one-half inch below the surface of the dirt. He controlled the temperature of the soil so it was above 60 degrees Fahrenheit. Nearly all the seeds sprouted.

He planted another batch of seeds two inches deep. He put them in a place that had a temperature that was less than 60 degrees. Almost none of those seeds sprouted.

Sam's conclusion

A temperature above 60 degrees causes the seeds to sprout.

Outline diagram

FIGURE 3 Sam's experiment.

logically support the opinion, and that the evidence must be accurate according to a reliable source. They learn to look for contradictions and inconsistencies, to look at all possible explanations for a set of facts, and then look for more information to rule out some of the explanations.

THE PROBLEM: AT-RISK SECONDARY SCHOOLS

Because the BIG Accommodation model improved the performance of students with disabilities on higher level cognitive tasks, we hypothesized that it would also work to accelerate the learning of large numbers of at-risk secondary students. Implementing the BIG Accommodation model in schools serving high-poverty neighborhoods would certainly present new problems that we had not yet encountered in our work with students with disabilities in middle class schools. Low achievement has been correlated with high poverty (Hodgkinson, 1992), English as a second language, and disability. These problems are difficult to resolve; breaking the cycle of low performance at the middle school level in high-poverty neighborhoods would indeed be challenging.

THE STRATEGY: BUILD A BIG BEACON

My approach was to work with a middle school that faced as many problems as possible to see what it would take to turn performance around. If performance could be turned around, then that site could be used as a training and demonstration center (a BIG Beacon) to teach and guide personnel from other schools with similar problems.

The school that fit those criteria was Charles M. Goethe Middle School in Sacramento, CA, the lowest performing middle school in one of the lowest performing districts in northern California at the time of our study. Goethe Middle School had a long-standing reputation of ineffectiveness. Of the elementary schools that fed into Goethe Middle School, most promoted a majority of students with mean percentile scores in the single digits. Ninety-five percent of the students received free or reduced-price lunches, 56% lived with families that received Aid for Dependent Children, approximately 40% were English language learners (ELL) from diverse language backgrounds, and 91% were comprised of various minorities.

GOETHE: THE MIDDLE SCHOOL THAT COULD

The transformation in the spirit of Goethe Middle School in the first year of the BIG Accommodation implementation was documented in a film (Palfremen, 1997) available from the Middle School Division of the Sacramento County Office of Education. The film includes testimonials and scenes describing the transformation of Goethe Middle School.

Quantitative evaluation of growth for low-performing middle school students has been problematic. Norm-referenced summaries of performance, such as the Stanford Achievement Test–Ninth Edition (SAT9), used for school accountability in California, are not sensitive to the growth of students at the low end of the distribution. To understand this lack of sensitivity, imagine a student performing at each percentile as a runner in a race of 100 competitors. The longer the race, the larger the gap: The main bunch of runners progress further down the road while stragglers spread out further behind. To pass 10 runners (or to gain 10 percentile points) requires 1 runner at the tail end of the pack to cover much more distance than 1 in the middle of the pack. Similarly, norm-referenced tests require more learning to gain points at the ends of the distribution than they require for students near the 50th percentile, or for groups who started later (younger students).

An example using real data from the Multilevel Academic Survey Test (MAST; Howell, Zucker, & Morehead, 1985), which has norms for groups in Grades 2 through 8, shows that the same amount of gain in raw score produces very different values for a student in Grade 3 versus one in Grade 8. For the same raw score gain from pretest (12) to posttest (28), a student in Grade 3 would show 48 percentile points gain (28 points on a normal curve equivalent [NCE] scale), whereas a student in Grade 8 would show only 4 percentile points gain (14 points on an NCE scale). This phenomenon occurs on any non-referenced test.

The MAST can be used to compare the performance of every student with that of norm groups from a wide range of age levels. For example, the raw score of 12 equaled the mean score of the second-grade group; the raw score of 28 equaled the mean score of the fourth-grade group. Both the Grade 3 and Grade 8 students in the example moved from working at a Grade-2 to a Grade-4 level. Both students made gains of 2 years in 1 year of academic work. This type of analysis represents the growth more fairly.

According to the MAST, the median-score student at Goethe increased two grade levels during the first year both in reading comprehension and in mathematics. The median score in reading improved from the fourth- to the sixth-grade level, and the median score in math improved from the fifth- to the seventh-grade level. In the analyses of the MAST scores from Goethe Middle school, no differences were found between seventh- and eighth-grade students; therefore, scores across both grades were aggregated for the analysis. Figure 4 shows how students at Goethe improved in grade-level equivalent scores in reading; Figure 5 shows how students improved in mathematics. The performance levels of the Grade-7 and Grade-8 students at Goethe Middle School shifted substantially from the lower to the higher grade-level ranges.

ELL

To determine the effect of the BIG Accommodation model on the performance of ELL, I analyzed pre and posttest scores of the MAST using the same cut scores. Figures 6 and 7 display the analyses for ELL. As the figures indicate, the shift from low to higher performance was stronger for ELL students in reading than for the group as a whole. The number of students reading at approximately grade level (Grade 7 and above) increased by more than 350%.

Results by Ethnic Group

Schoolwide data from the MAST were disaggregated by ethnic group. The results of this analysis for Year 1 for reading are displayed in Table 3. In reading, all ethnic groups made substantial progress. The mathematics instruction had more differential effects by ethnic group. Perhaps the different grouping arrangements used for math and reading can explain these differences. For reading instruction, students were mixed ethnically and grouped according to specific reading needs. For mathematics instruction, students were grouped according to their English language level, with native speakers grouped separately from ELL, who were taught in

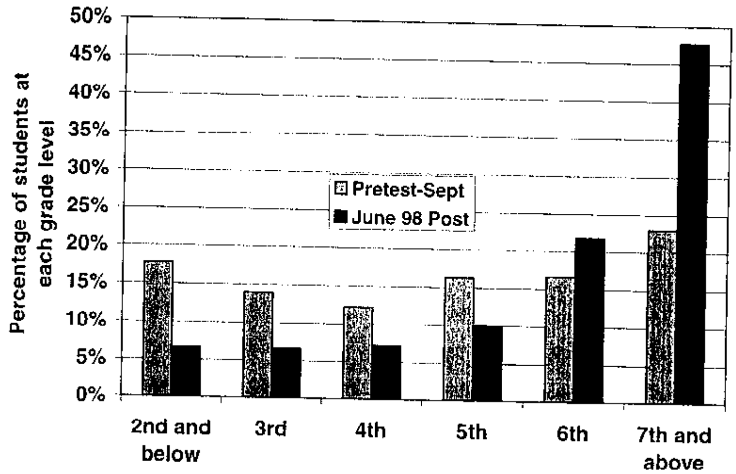


FIGURE 4 Year 1 change in reading performance (students performing at each grade level) on the Multilevel Academic Survey Test, September 1997 to June 1998 ($n = 518$).

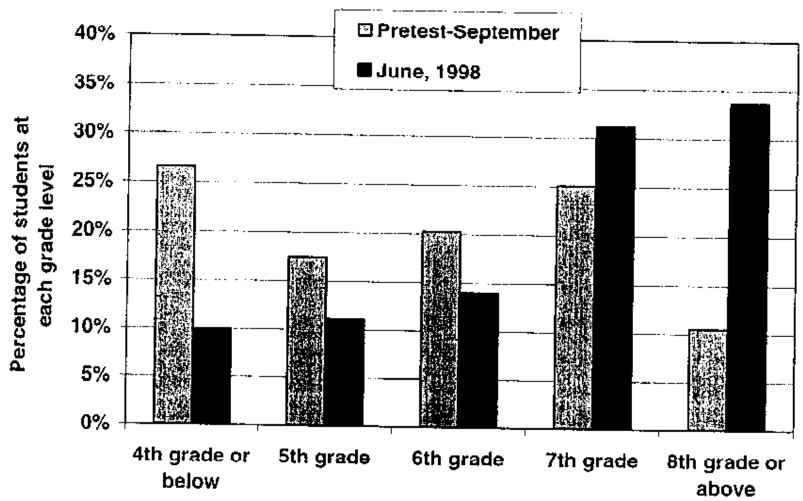


FIGURE 5 Year 1 change in mathematics performance (students performing at each grade level) on the Multilevel Academic Survey Test, September 1997 to June 1998 ($n = 518$).

smaller groups (20). In effect, this meant that large numbers of African Americans and very few White English-speaking students were taught in groups of 30 to 35. (In Year 2, students were grouped for math instruction according to their skill needs, regardless of their language level or disability. Consequently, results were consistent in mathematics across all ethnic groups.)

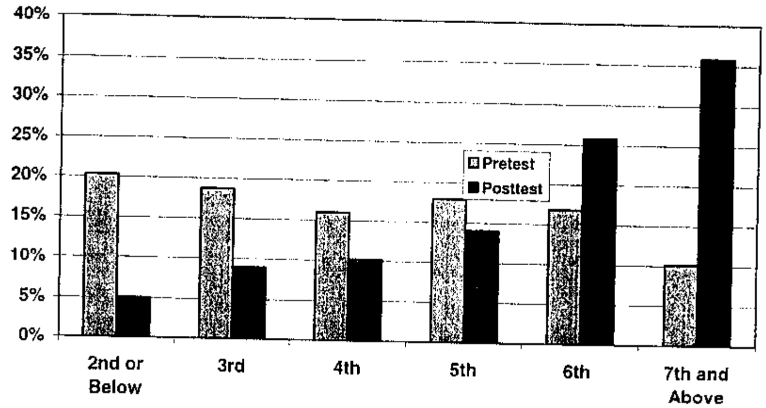


FIGURE 6 Year 1 change in reading performance (students performing at each grade level) for English language learners on the Multilevel Academic Survey Test, September 1997 to June 1998 ($n = 256$).

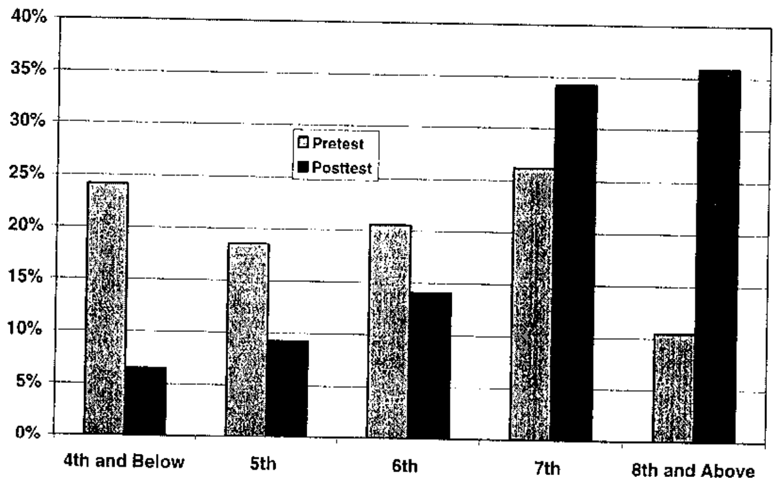


FIGURE 7 Year 1 change in mathematics performance (students performing at each grade level) for English language learners on the Multilevel Academic Survey Test, September 1997 to June 1998 ($n = 256$).

Gifted Students

Gifted students were also placed in the high-level reading program (*Reasoning and Writing: Level F,* Engelmann & Grossen, 2001b). Table 4 shows the percentile gains made by the Grade-7 and Grade-8 gifted groups on the MAST.

TABLE 3
Grade Equivalent Gains on MAST in Reading, by Ethnic Group

Ethnicity	No.	Grade-Level Equivalent Pre	Grade-Level Equivalent Post	Change
Asian	210	4.6	5.9	+1.3
Black	142	4.6	5.8	+1.2
Latino	120	3.9	5.5	+1.6
White	48	5.2	7.3	+2.1
Native American	4	3.2	4.9	+1.7

Reduced Behavior Problems

During the first year of implementation, the last period of the school day was the only period in which students in all classrooms were being taught using DI. During the other periods of the day, the instruction was mixed; some groups were learning with DI, others were not. The vice principal reported that during the last period of the day, there were rarely any referrals for behavior. During each of the earlier periods of the day, there was an average of 8 to10 referrals per class period. The success students were experiencing in learning seemed to be reflected in their behavior.

YEAR 2: GOETHE MIDDLE SCHOOL BECOMES A BIG BEACON

California's statewide assessment using the SAT9 began in spring 1998. Because performance on the SAT9 has been the focus for accountability and evaluation, I used primarily the SAT9 for Year 2 of the DI implementation at Goethe Middle School. Being a statewide measure, the SAT9 also allowed comparisons with other schools. Concurrently, I continued to use the MAST for subgroup analyses. The subgroup analyses for Year 2 with the MAST were similar to the Year 1 reports, so they are not reported here. The correlation of the MAST raw scores with the SAT9 percentile scores was quite high (Pearson $r = .76$).

Longitudinal Comparisons

The state of California publishes percentile equivalents for mean raw scores for all grade levels at all schools in California. The published mean scores in percentiles for the same cohort for reading were, for Grade 7 (1998), $M = 21$st percentile; for Grade 8 (1999), $M = 35$th percentile. According to the MAST the progress during Year 1 for the same cohort in reading was a change from a median percentile of 9 to a median percentile of 25. Although 35th percentile at the

TABLE 4
Percentile Gains in Multilevel Academic Survey Test Reading for Gifted Students

Grade	*Percentile for the Mean Pretest Score*	*Percentile for the Mean Posttest Score*
7	35	86
8	36	85

end of Year 2 is still well below the mean, the extent of the improvement after implementing DI seems substantial.

For mathematics, the same cohort scored as follows on the SAT9: For Grade 7 (1998), 27 = mean percentile; for Grade 8 (1999), 35 = mean percentile.

Comparisons

The SAT9 also allowed comparisons with other schools. A comparison school was selected in the same district. The comparison school had very similar demographic statistics, the same pretest score on the SAT9, and was also implementing a new research-based reading program (although, not the BIG Accommodation model). The scores for the same cohort at the comparison school were as follows: In Grade 7 (1998), the score was 21; in Grade 8 (1999), the score was 29.

Officials in the California Department of Education reviewed all the middle schools in the state and found that the gains that Goethe showed in reading for the same cohort ranked fifth highest in the state. The schools making greater gains than Goethe Middle School were all schools with higher initial pretest performance. No low-achieving school made gains comparable to those made at Goethe. Generally, it has been difficult to turn around low-achieving middle schools.

By Level

Raw scores on the MAST at the low end of the distribution had a lower correlation with the SAT9 percentile scores than MAST raw scores at the high end of the distribution. This finding corroborated a theory that the MAST raw scores were more sensitive to gains for very low-achieving students, whereas the SAT9 was more sensitive to the learning gains of higher performing students.

On the MAST, students at all instructional levels made 2 years gain for 1 year of instruction. The SAT9, however, seemed to show little gain for the lower instructional levels and higher gain for the higher levels.

To see which reading program levels in the DI model contributed most to the SAT9 gains, I used a "net shift analysis" (Lindman, 1968). Although about as dated as *t* tests, the net shift analysis provided a means for weighting the low- and high-end scores of the distribution to compensate for the problems in percentile comparisons described earlier. With the net shift analysis, individual percentile score distributions can be used to evaluate growth, on a relative scale, by comparing distributions of test scores made by groups of pupils on standard tests with distributions made by other groups of students on the same tests.

By identifying the percentage of student scores that must be shifted to an adjacent cell (interval) to make the two distributions equivalent, the net shift analysis technique reveals changes in score distributions that may occur when different teaching methods are used. The technique also offers a more complete comparison between the distribution of scores made by a selected group of pupils and a norm group. Table 5 displays the net shift analysis values for each of the program levels used for reading at Goethe, including both the weighting to adjust low scores and the weighting to adjust high scores.

Regardless of whether the low- or the high-end scores are weighted using the net shift analysis, the higher levels of the programs show greater SAT9 gains than the low-end programs. The content of the higher level programs aligns better with the content of the SAT9 and with the grade-level standards. However, the lower level programs build the necessary foundation for students to be successful in the higher level programs. Clearly, it is necessary to implement the higher level programs to achieve significant gains on the SAT9.

YEAR 3: RESOURCE-EFFICIENT DISSEMINATION

Because of the improvement in performance at Goethe, and the growing awareness in all of California of the extent of illiteracy in secondary schools, both middle and high schools were eager to implement the decoding program used in the BIG Accommodation, the Corrective Reading program (Engelmann et al., 1999). Consequently, we turned our attention to developing an efficient model for training and replication.

Early Intensive Coaching

Designing lessons that are successful with diverse learners (because they incorporate the six principles of accommodation) is significantly more time consuming than actually delivering the lessons. Teachers simply do not have the time to design the lessons themselves. Consequently, DI models give the teachers the core lesson plans for teaching the big ideas. By having ready-made lesson plans, training can proceed more efficiently.

TABLE 5
Net Shift Analysis for Each Reading Program Level at Goethe Middle School

Net Shift	*Corrective reading, Decoding, Level B1*	*Corrective reading, Decoding, Level B2*	*Corrective reading, Decoding, Level C*	*Corrective reading, Comprehension, Level C*	*Reasoning and writing, Level E*	*Reasoning and writing, Level F*
n	22	47	33	42	149	124
Weighted low scores	8%	10%	36%	34%	34%	52%
Weighted high scores	9%	11%	42%	45%	45%	72%

No matter how efficient the training might be, it seems that in-class coaching remains a necessity. Joyce and Showers (1995) documented the extent to which different types of teacher training resulted in teachers applying what they learn in the classroom. They found that professional development that includes demonstration and simulated practice transfers to the classroom with 10% to 15% fidelity. With the addition of coaching, the rate of transfer increases to between 80% and 90%.

Because the DI procedures are so specific, workshop training often leaves teachers with the impression that a good implementation merely follows these procedures in a rather "robot-like" fashion. This is not the case. A teacher who implements well is constantly responding to nuances in the students' performance. Learning this kind of responsiveness in simulated practice conditions with adults is almost impossible, because when proficient persons role-play naive learners they have a great deal of difficulty responding the same way naive learners do.

The skills involved in bringing the low performers to criterion, while keeping the lessons challenging for the most proficient, seem best taught to teacher trainees in the context of watching a model with real students in the classroom, while having an opportunity to copy the model. In the resource-efficient training system I developed for replicating the knowledge gained at Goethe, model schools host training sessions where teachers in training actually practice the presentation techniques with master teachers in the model school with the master teachers' students. The master teacher first models the techniques, then team-teaches with the trainee, and finally monitors and gives feedback to the teacher who by that stage should have become quite proficient in the strategies. With this training model, teachers are far more likely to use the programs in a manner that maximizes gains in their respective classrooms.

Even after teachers return to their classrooms, follow-up visits by an expert guest coach are still necessary, but do not seem to be required nearly to the extent that they are required when only workshop training is provided. Exper guest coaches work with teachers in their classrooms until the teachers and their instructional groups achieve the following four criteria:

1. Students respond correctly to new items that incorporate previously taught concepts 90% of the time.
2. Students are 100% on task 100% of the time.
3. Teachers give feedback to students in a ratio of three statements of specific praise for each corrective statement.
4. Students display pride and joy in their work.

Coaches are trained in specific techniques for helping teachers achieve the aforementioned goals and provide this coaching support for teachers in an in-class format. In-class coaching results in faster acquisition of new teaching behaviors than after-class feedback (Coulter, 1997). In this model, the coach identifies a specific teacher behavior aimed at improving student performance, models this behavior while the trainee is leading the teaching, observes the teacher incorporating the modeled behavior, and checks if the change in teaching behavior is yielding the expected change in student performance. (For more information about the specific coaching and training model, see Grossen & Scott, 2000.)

Electronic Progress Monitoring

In addition to the use of the BIG Accommodation curricula ("B") and in-class coaching ("I"), a third important factor of success is great expectations ("G"). Schools achieve great expectations with the assistance of an electronic system of continuous progress monitoring (Caros, 2001). The statewide accountability tests that occur at the end of the year come too late for anybody to identify problems and solve them. The progress monitoring system provides a formative evaluation process, which works much like a physicians' stethoscope, monitoring the health of the implementation on an ongoing basis. Teachers routinely enter their student mastery test results for each instructional group into a standard data entry form. Every 4 weeks, data showing the percentage of student at mastery in each instructional group and the rate of progress through the lessons are summarized and reported.

One report provides information about the rate of progress of all the instructional groups through the lessons as well as the percentage of students who are at mastery. A second report provides a list of the students who are not at mastery along with a possible reason, identified by the teacher. These reports serve as an efficient means for preventing eventual failure by guiding the focus for immediate coaching on the specific groups and individual students needing help, and also serve as a powerful means of motivating students and teachers by routinely recognizing their achievement and progress.

The goal of BIG Accommodation is for all students to be successful on all parts of the mastery tests at all times. Students who are proficient every step of the way are more likely to be proficient on statewide assessments. The coordinated efforts of all support personnel are required to support this goal.

SWAT teams (teams for providing extra support to the "students who wannabe achievers too") are organized to respond to the information in these reports. For example, one replication site scheduled a first-period schoolwide reading block. In the early progress reports, teachers identified attendance and tardiness problems as the primary reason for students who had not achieved mastery. The SWAT team involved the parent council in seeking a solution to the problem. The parent council provided funds to purchase incentives for students present at the first minute of school. Subsequent progress reports allowed the parent council to assess whether the investment was worthwhile. It was; mastery increased substantially.

Figure 8 illustrates how the data are used in the schoolwide system of support for teachers, which is also initially supported by external coaches who work with both the teachers and the support staff in providing research-based solutions to problems. The focus of all personnel is on enhancing the instructional interaction between teachers and students.

Results of the Replication Efforts of the Reading Component of the BIG Accommodation Model

In Year 3, I established two more training centers, one at Apple Valley Middle School in San Bernadino County and one at Raymond Cree Middle School in Riverside County. Subsequently, both schools more than doubled their target growth rate for the year and were placed on the list to receive the largest monetary awards from the state of California, Certificated Staff Performance Incentives, set originally at up to $25,000 per teacher.

Other training centers were established in northern California through the Sacramento County Office of Education at the following locations: Starr King Middle School, LeRoy Greene Middle School in Sacramento County, and at South Lake Tahoe Middle School. Over 800 teachers in over 50 schools in more than 30 districts were trained in the remedial reading programs. Fewer schools implemented the higher level reading programs.

Of these schools, several used the MAST to evaluate the performance of students placed in the decoding levels of the Corrective Reading program (Engelmann et al., 1999). Table 6 displays a summary of these results for all 11 schools. Instructional methods that are less accommodating of diversity did not raise the performance of students in the lowest ranges of performance, as seen in

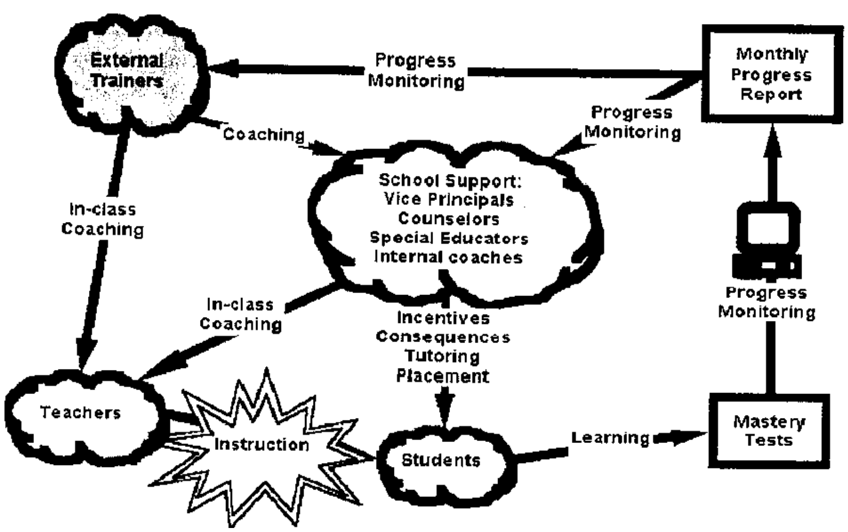


TABLE 6
Change in Literacy Levels of Students Placed in Corrective Reading—Decoding for 11 Schools in California

		Percentage Scoring Below 2nd-Grade Level			*Percentage Scoring Above 5th-Grade Level*		
Schools	*No.*	*Pre*	*Post*	*Change*	*Pre*	*Post*	*Change*
Control school	59	56	53	−3	5	14	+9
S1[a]	245	55	31	−24	9	24	+15
S2[b]	59	44	25	−19	10	37	+27
S3[a]	129	44	18	−26	13	33	+20
S4[b]	150	41	15	−26	16	38	+22
S5[a]	282	34	14	−20	20	53	+33
S6[b]	338	34	14	−20	22	55	+33
S7[a]	183	27	13	−14	6	21	+15
S8[b]	82	24	13	−11	20	43	+23
S9[b,c]	110	23	6	−17	20	49	+29
S10[b,c]	558	17	9	−8	42	56	+14
S11[b]	455	14	5	−9	44	61	+17

[a]The period between pre and posttest for these schools is 4 months. [b]The period between pre and posttest for these schools is 8 months. [c]These schools are in their second year of implementation.

the control school in Table 6. Only 3% of the students in the control sample moved out of the second-grade-and-below performance level, where 56% of the sample performed. Schools implementing Corrective Reading (Engelmann et al.), an instructional model designed specifically to raise the performance of all students, were able to move substantial numbers of students out of the lowest performance ranges. Even schools beginning implementation only in the second semester made substantial progress with the students in the lowest performance ranges by the end of the year. Schools in the second year of implementation reduced the percentage of students performing in the lowest range to less than 10% of the sample. These two schools, one of them Goethe Middle School, were both schools with demographic profiles representative of the most troubled schools in America.[1]

External coaches identified schools that implemented components of the BIG Accommodation with fidelity, using an implementation fidelity rubric. These schools were categorized according to whether they implemented only the low-level programs for decoding, or whether they also included an implementation of the higher level reading programs. Table 7 displays these categories in

[1]Color pie charts displaying the results reported in Table 6 in greater detail can be found at www.higherscores.org.

TABLE 7
Summary of SAT9 Performance for BIG Schools, Compared to California Eighth Graders

Grade 8 SAT9 Scores	*Number of Schools*	*Mean SAT9 Percentile[a]: Pre (1999)*	*Mean SAT9 Percentile[a]: Post (2000)*	*Change in Mean*	*Percentage of Schools in Quartile 1: Pre (1999)*	*Percentage of Schools in Quartile 1: Post (2000)*	*Percentage up to Quartile 2*
Full BIG[b]	1	21	35	14	58	39	19
Schoolwide BIG reading[c]	7	32	40	8	47	37	10
Partial BIG[d]	21	49	55	6	27	20	7
California eighth graders[e]		43	47	4	33	27	6

Note. SAT9 = Stanford Achievement test–Ninth Edition.

[a]Percentiles of the mean raw scores provided on the California Department of Education web site for each school. [b]Full BIG: All programs implemented, including *Understanding U.S. History* (Carnine, Crawford, Harniss, & Hollenbeck, 1998) and *Expressive Writing 2* (Engelmann & Silbert, 1983). [c]Schoolwide BIG Reading: Schools tested all students for placement in a reading program, Corrective Reading, (Engelmann et al., 1999); "decoding" and "comprehension") and Reasoning and Writing (Engelmann & Grossen, 2001a, 2001b), as in the Goethe Project. [d]Partial BIG: Schools selected students scoring below the 25th or 35th percentile in reading, and placed only those students in corrective reading (Engelmann et al., 1999; "decoding"), the remedial component of BIG. [e]California: Percentiles for the mean raw scores of all the Grade-8 students in California (www.cde.ca.gov).

the left column. For each of these categories of schools, the percentages of students performing in each quartile of the SAT9 were collected from the California state Web page (www.cde.ca.gov). Table 7 displays the changes in mean scores and the movement out of the bottom quartile for these schools.

No schools have replicated the entire BIG Accommodation model. The bigger gain in performance for schools implementing the higher level reading programs in addition to the lower level programs was consistent with the earlier analysis that the higher level programs contribute more to gains on standardized test scores, although learning gains for the lower level programs are important and significant.

DISCUSSION: THE WORK CONTINUES

Two significant achievements of the Goethe Research Project were the development of a training model that results in consistently high levels of success across implementations, and evidence that the newer generation DI programs have been very effective in significantly raising the performance of students performing at or near grade level.

Implementing the BIG Accommodation with fidelity in low-achieving schools is a challenging task. The concept of building a BIG Beacon school, which provides a model of a high-quality implementation and serves as a training site, seems clearly the most resource-efficient way to expand. I am working to develop BIG Beacons in Florida, Kansas, and Hawaii. The work continues in California as well.[2]

REFERENCES

Carnine, D. (1994). The BIG Accommodation program. *Educational Leadership 51*(6), 87–88.

Carnine, D., Caros, J., Crawford, D., Hollenbeck, K., & Harniss, M. (1996). Designing effective U.S. history curricula for all students. In J. Brophy (Ed.), *Sixth handbook on research on teaching: Vol. 6. Advances in research on teaching* (pp. 207–256). Greenwich, CT: JAI.

Carnine, D., Crawford, D., Harniss, M., & Hollenbeck, K. (1998). *Understanding U.S. history.* Eugene: University of Oregon.

Caros, J. (2001). *Emerald ice: Fourth-dimension database.* Vancouver, BC: Seira.

Collins, M., & Carnine, D. (1988). Evaluating the field test revision process by comparing two versions of a reasoning skills CAI program. *Journal of Learning Disabilities, 21,* 375–379.

Coulter, G. (1997). *An examination of the effectiveness of in-class instructive feedback and afterclass instructive feedback for teachers learning specific teaching behaviors.* Eugene: University of Oregon.

Dixon, R., & Engelmann, S. (1999). *Spelling through morphographs.* Blacklick, OH: Science Research Associates.

[2]Further information on the available trainings can be found at www.higherscores.org.

Engelmann, S., Carnine, L., Johnson, G., Hanner, S., Meyer, L., & Osborne, S. (1999). *Corrective reading.* Blacklick, OH: Science Research Associates.

Engelmann, S., & Grossen, B. (2001a). *Reasoning and writing: Level E.* Blacklick, OH: Science Research Associates.

Engelmann, S., & Grossen, B. (2001b). *Reasoning and writing: Level F.* Blacklick, OH: Science Research Associates.

Engelmann, S., & Silbert, J. (1983). *Expressive Writing 2.* Blacklick, OH: Science Research Associates.

Grossen, B. (1998). *The research base for corrective reading: SRA.* Blacklick, OH: Science Research Associates.

Grossen, B., & Carnine, D. (1990). Diagraming a logic strategy: Effects on more difficult problem types and transfer. *Learning Disability Quarterly, 13,* 168–182.

Grossen, B., Carnine, D., & Lee, C. (1996). *The effects of instruction designed for diverse learners and constructivist instruction on middle-school students' achievement and problem solving in earth science.* Unpublished manuscript, University of Oregon, Eugene.

Grossen, B., & Ewing, S. (1996). Raising mathematics problem-solving performance: Do the NCTM teaching standards help? (Final report). *Effective School Practices, 13*(2), 79–91.

Grossen, B., & Scott, C. (2000). *A case study of a professional development center model for disseminating direct instruction.* Unpublished manuscript, University of Oregon, Eugene.

Hodgkinson, H. L. (1992). *A demographic look at tomorrow.* Washington, DC: Institute for Educational Leadership, Center for Demographic Policy.

Hofmeister, A., Engelmann, S., & Carnine, D. (1989). Developing and validating science education videodisks. *Journal of Research in Science Teaching, 26,* 665–677.

Howell, K., Zucker, S., & Morehead, M. (1985). *MAST: Multilevel Academic Survey Test.* San Antonio, TX: Psychological Corporation.

Joyce, B., & Showers, B. (1995). *Student achievement through staff development: Fundamentals of school renewal.* New York: Longman.

Kameenui, E., & Carnine, D. (2001). *Effective teaching strategies that accommodate diverse learners.* Columbus, OH: Merrill.

Lindman, E. (1968). *Net-shift analysis for comparing distributions of test scores.* Los Angeles, CA: University of California, Center for the Study of Evaluation of Instructional Programs.

Moore, L., & Carnine, D. (1989). Evaluating curriculum design in the context of active teaching. *Remedial and Special Education, 10*(4), 28–37.

Niedelman, M. (1992). Problem solving and transfer. In D. Carnine & E. Kameenui (Eds.), *Higher order thinking: Designing curriculum for mainstreamed students* (pp. 137–156). Austin, TX: PRO-ED.

Palfremen, J. (1997). *Goethe: The middle school that could* [Videotape]. Boston: Palfremen Film Group.

Schneps, M. H. (Director/Producer). (1987). *A private universe* [Motion picture]. Boston: Harvard University and Smithsonian Institute.

Systems Impact Incorporated. (1987). *Earth science.* St. Louis, MO: Phoenix Film.

Woodward, J. (1994). Effects of curriculum discourse style on eighth graders' recall and problem solving in earth science. *The Elementary School Journal, 94,* 299–314.

Woodward, J., Carnine, D., & Gersten, R. (1988). Teaching problem solving through computer simulations. *American Educational Research Journal, 25*(1), 72–86.

Woodward, J., & Noell, J. (1992). Science instruction at the secondary level: Implications for students with learning disabilities. In D. Carnine & E. Kameenui (Eds.), *Higher order thinking: Designing curriculum for mainstreamed students* (pp. 39–58). Austin, TX: PRO-ED.