# The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring

BENJAMIN S. BLOOM
*University of Chicago and Northwestern University*

Two University of Chicago doctoral students in education, Anania (1982, 1983) and Burke (1984), completed dissertations in which they compared student learning under the following three conditions of instruction:

1. *Conventional.* Students learn the subject matter in a class with about 30 students per teacher. Tests are given periodically for marking the students.
2. *Mastery Learning.* Students learn the subject matter in a class with about 30 students per teacher. The instruction is the same as in the conventional class (usually with the same teacher). Formative tests (the same tests used with the conventional group) are given for feedback followed by corrective procedures and parallel formative tests to determine the extent to which the students have mastered the subject matter.
3. *Tutoring.* Students learn the subject matter with a good tutor for each student (or for two or three students simultaneously). This tutoring instruction is followed periodically by formative tests, feedback-corrective procedures, and parallel formative tests as in the mastery learning classes. It should be pointed out that the need for corrective work under tutoring is very small.

*Benjamin S. Bloom is Professor of Education, University of Chicago and Northwestern University, 5835 S. Kimbark Avenue, Chicago, Illinois 60637. His areas of specialization are evaluation and the teaching-learning process.*

The students were randomly assigned the three learning conditions, and their initial aptitude tests scores, previous achievement in the subject, and initial attitudes and interests in the subject were similar. The amount of time for instruction was the same in all three groups except for the corrective work in the mastery learning and tutoring groups. Burke (1984) and Anania (1982, 1983) replicated the study with four different samples of students at grades four, five, and eight and with two different subject matters, Probability and Cartography. In each sub-study, the instructional treatment was limited to 11 periods of instruction over a 3-week block of time.

Most striking were the differences in final achievement measures under the three conditions. Using the standard deviation (sigma) of the control (conventional) class, it was typically found that the average student under tutoring was about two standard deviations above the average of the control class (the average tutored student was above 98% of the students in the control class).[1] The average student under mastery learning was about one standard deviation above the average of the control class (the average mastery learning student was above 84% of the students in the control class).

The variation of the students' achievement also changed under these learning conditions such that about 90% of the tutored students and 70% of the mastery learning students attained the level of summative achievement reached by only the highest 20% of the students under conventional instructional conditions. (See Figure 1.)

There were corresponding changes in students' time on task in the classroom (65% under conventional instruction, 75% under Mastery Learning, and 90+% under tutoring) and students' attitudes and interests (least positive under conventional instruction and most positive under tutoring). There were great reductions in the relations between prior measures (aptitude or achievement) and the summative achievement measures. Typically, the aptitude-achievement correlations changed from +.60 under conventional to +.35 under mastery learning and +.25 under tutoring. It is recognized that the correlations for the mastery learning and tutoring groups were so low because of the restricted range of scores under these learning conditions. However, the most striking of the findings is that under the best learning conditions we can devise (tutoring), the average student is 2 sigma above the average control student taught under conventional group methods of instruction.

The tutoring process demonstrates that *most* of the students do have the potential to reach this high level of learning. I believe an important task of research and instruction is to seek ways of accomplishing this under more practical and realistic conditions than the one-to-one tutoring, which is too costly for most societies to bear on a large scale. This is the "2 sigma" problem. Can researchers and teachers devise teaching-learning conditions that will enable the majority of students under *group instruction* to attain levels of achievement that can at present be reached only under good tutoring conditions?

It has taken almost a decade and a half to develop the Mastery Learning (ML) strategy to a point where large numbers of teachers at every level of instruction and in many countries can use the feed-back-corrective procedures to get the 1 sigma effect (the average ML student is above 84% of the students under conventional instruction—even with the same teacher teaching both the ML and the conventional classes). If the research on the 2 sigma problem yields *practical methods* (methods that the average teacher or school faculty can learn in a brief period of time and use with little more cost or time than conventional instruction), it would be an educational contribution of the greatest magnitude. It would change popular notions about human potential and would have significant effects on what the schools can and should do with the educational years each society requires of its young people.

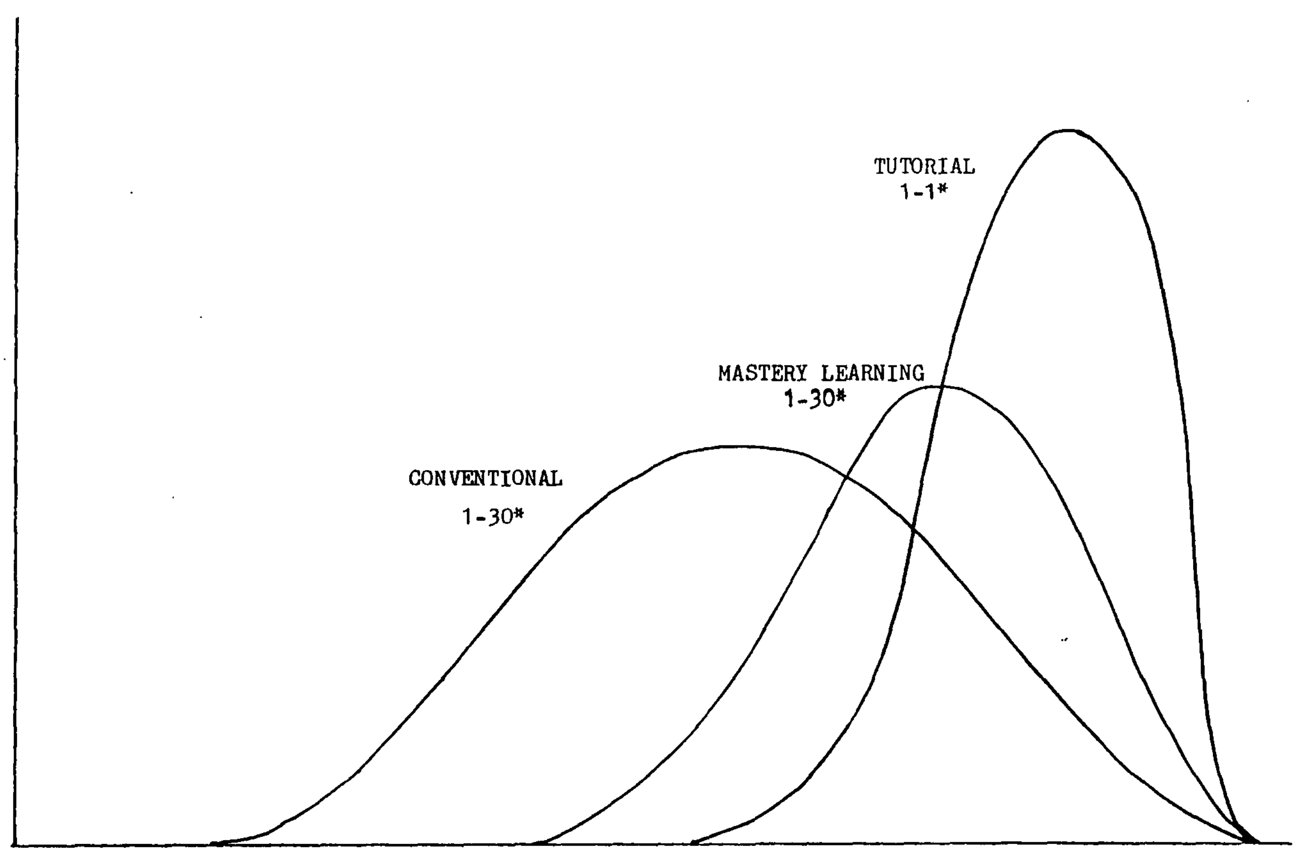
This paper is a brief presentation of the work on solutions to the 2 sigma problem. It is hoped that it will interest both educational researchers and teachers in further research and application of these ideas.

## The Search

In a number of articles, my graduate students and I have attempted to contrast alterable educational variables with more stable or static variables (Bloom, 1980). In our treatment of this topic, we summarized the literature on such alterable variables as the *quality of teaching*, the *use of time* by teachers and students, *cognitive* and *affective* entry characteristics of students, *formative testing, rate of learning*, and the *home environment*. In each case, we contrasted these alterable variables with the more *stable* variables (e.g., personal characteristics of teachers, intelligence measures, achievement tests for grading purposes, socioeconomic status of the family, etc.) and indicated some of the ways in which the alterable variables influence learning and the processes by which these variables have been altered.

But not all alterable variables are likely to have equal effects on learning. Our research summaries were intended to emphasize the alterable variables that have had the strongest effects on school learning. Within the last 3 years, this search has been aided by the rapid growth of the meta-analysis literature. In this literature, each writer has summarized the research literature on a particular set of alterable variables to indicate the effect size between control and experimental groups of students. They have standardized

**FIGURE 1.** *Achievement distribution for students under conventional, mastery learning, and tutorial instruction.*




*Teacher-student ratio

the results in terms of the *difference* between the experimental and control groups divided by the standard deviation of the control group.[2]

In each study, the reviewer also analyzed the effect size under different conditions, level of school, sex of student, school subject, size of sample, and so on. Such reviews are very useful in selecting alterable variables that are most likely to contribute significantly to the 2 sigma solution.

Table I is adapted from a summary of effect sizes of key variables by Walberg (1984) who, with other co-authors, has contributed greatly to this literature. In Table I he has listed the selected variables in order of magnitude of effect size. (We have added other variables and indicated the equivalent percentile for each effect size.) Thus, in the first entry, *tutorial instruction,* we have indicated the effect size (2 sigma) and indicated that under tutorial instruction, the average student is above 98% of the students under the control teaching conditions. A list of effect size studies appears in the Appendix at the end of this article.

In our own attempts to solve the 2 sigma problem we assume that two or three alterable variables must be used that *together* contribute more to the learning than any one of them alone. Because of more than 15 years of experience with ML at different levels of education and in different countries, we have come to rely on ML as one of the possible variables to be combined with selected other variables. ML (the feedback-corrective process) under good conditions yields approximately a 1 sigma effect size. We have systematically tried other variables which, in combination with ML, might approach the 2 sigma effect size. So far, we have *not* found any two variable combination that has exceeded the 2 sigma effect. Thus, some of our present research reaches the 2 sigma effect but does not go beyond it.

We have classified the variables in Table I in terms of the direct object of the change process: (a) the learner; (b) the instructional material; (c) the home environment or peer group; and (d) the teacher and the teaching process.

We have speculated that two variables involving different objects of the change process may, in some instances, be additive, whereas two variables involving the same object of the change process are less likely to be additive (unless they occur at different times in the teaching-learning process). Our research is intended to determine when these rules are true and when they are not. Several of the studies done so far suggest that they may be true. Thus the ML process (which affects the learner most directly), when combined with changes in the teaching process (which affects the teacher most directly), yield additive results. (See Tenenbaum, p. 13 of this article and Mevarech, p. 14 of this article). Although we do not believe these two rules are more than *suggestive* at present, future research on this problem will undoubtedly yield a stronger set of generalizations about how the effects of separable variables may be best combined.

In our work so far we have restricted the search to two or three variables, each of which is likely to have a .5 sigma effect or greater. We suspect that the research, as well as the applications to school situations, would get too complex if more than three alterable variables are used. In any case, our work has begun with variables in the top half of Table I. Perhaps as the research moves on, it will be necessary to include some of the variables in the lower part of Table I.

In our research with two variables, we have made use of a 2 × 2 randomized design with ML and one other variable. So far we have not done research with three variables. Where possible, we try to replicate the study with at least two subject fields, two levels of schooling, or some combination of subject fields and levels of schooling. We hope that others will take up this 2 sigma search and that some guide-

TABLE I

*Effect of selected alterable variables on student achievement (see Appendix)*

| | | Effect size | Percentile equivalent |
|---|---|---|---|
| D[a] | Tutorial instruction | 2.00 | 98 |
| D | Reinforcement | 1.20 | |
| A | Feedback-corrective (ML) | 1.00 | 84 |
| D | Cues and explanations | 1.00 | |
| (A)D | Student classroom participation | 1.00 | |
| A | Student time on task | 1.00[b] | |
| A | Improved reading/study skills | 1.00 | |
| C | Cooperative learning | .80 | 79 |
| D | Homework (graded) | .80 | |
| D | Classroom morale | .60 | 73 |
| A | Initial cognitive prerequisites | .60 | |
| C | Home environment intervention | .50[b] | 69 |
| D | Peer and cross-age remedial tutoring | .40 | 66 |
| D | Homework (assigned) | .30 | 62 |
| D | Higher order questions | .30 | |
| (D)B | New science & math curricula | .30[b] | |
| D | Teacher expectancy | .30 | |
| C | Peer group influence | .20 | 58 |
| B | Advance organizers | .20 | |
| | Socio-economic status (for contrast) | .25 | 60 |

*Note.* This table was adapted from Walberg (1984) by Bloom.

[a]*Object of change process*—A-Learner; B-Instructional Material; C-Home environment or peer group; D-Teacher.

[b]Averaged or estimated from correlational data or from several effect sizes.

lines for the research can be set up to make the combined results more useful and to reduce the time and costs for experimental and demonstration studies.

### Improving Student Processing of Conventional Instruction

In this section of the paper we are concerned with ways in which students can learn more effectively without basically changing the teaching. If students develop good study habits, devote more time to the learning, improve their reading skills, and so on, they will be better able to learn from a particular teacher and course—even though neither the course nor the teacher has undergone a change process.

For example, the ML feedback-corrective approach is addressed primarily to providing students with the cognitive and affective prerequisites for each new learning task. As we have noted before, when the ML procedures are done systematically and well, the school achievement of the average student under ML is approximately 1 sigma (84 percentile) above the average student in the control class, even when both classes are taught by the *same teacher* with much the same instruction and instructional material. As we view the ML process, we regard it as a method of improving the students' learning from the same teaching over a series of learning tasks.

The major changes under the ML process are that more of the students have the cognitive prerequisites for each new learning task, they become more positive about their ability to learn the subject, and they put in more active learning time than do the control students. As we observe the students' learning and the test results in the ML and the conventional class, we note the improvements in the student learning under ML and the lack of such improvement in conventional classes.

One of our University of Chicago doctoral students, Leyton (1983), suggested that one approach to the 2 sigma problem would be to use ML during the advanced course in a sequence, but in addition attempt to *enhance the students' initial cognitive entry prerequisites* at the beginning of the course. Working with high school teachers in Algebra 2 and French 2, they developed an initial test of the prerequisites for each of these courses. The procedure in developing the initial test was to take the final examination in the prior course (Algebra 1 or French 1) and have a committee of four to six teachers in the subject independently check each test item that they believed measured an idea or skill that was a necessary prerequisite for the next course in the subject. There was very high agreement on most of the selected items, and discussion among the teachers led to consensus about some of the remaining items.

Two of the classes were helped to review and relearn the specific prerequisites they lacked. This was not done for the students in the other two classes—they spent the time on a more general and informal review of the content taught in the previous course (Algebra 1 or French 1). The method of enhancing the prerequisites was much like the ML feedback-corrective process where the teacher retaught the items that the majority of students had missed, small groups of students helped each other over items that had been missed, and the students reviewed items they were not sure about by referring to the designated pages in the instructional material. The corrective process took about 3 to 4 hours during the first week of the course. After the students completed the corrective process, they were given a parallel test. As a result of the corrective process, most of the students reached the mastery standard (80%) on the parallel test given at the end of the first week of the course. In a few cases, students who didn't reach this standard were given further help.

More important was the improved performance of the enhanced classes over the other two classes on the first *formative* test in the advanced course (French 2 or Algebra 2). The two enhanced classes, which had been helped on the initial prerequisites, were approximately .7 sigma higher than the other two classes on the first formative test given at the end of a 2-week period of learning in the advanced course.

When one of the enhanced classes was also provided with ML feedback-corrective procedures over a series of learning tasks, the final results after a 10- to 12-week period of instruction was that this experimental group was approximately 1.6 sigma above the control group on the summative examination. (The average student in the ML plus enhanced initial prerequisites was above 95% of the control students on this examination.) There were also attitudinal and other affective differences in students related to these achievement differences. These included positive academic self-concept, greater interest in the subject, and greater desire to learn more in the subject field.

In Leyton's (1983) study, he found that the average effect of initial enhancement of prerequisites alone is about .6 sigma (see differences between conventional and conventional plus enhanced prerequisites and between ML and ML plus enhanced prerequisites in Figure 2). That is, we have two processes — *ML* and *initial enhancement of cognitive prerequisites* — that have sizeable but separate effects. When they are combined, their separate effects tend to be additive. We believe these two variables are additive because they occur at different times. The enhancement of the initial prerequisites is completed during the first week of the new course, while the ML feedback-corrective process takes place every 2 or 3 weeks during the course, after the initial enhancement.

This solution to the 2 sigma problem is likely to be applicable to sequential courses in most school subjects. (In the United States, over two-thirds of the academic courses in elementary-secondary schools are sequential courses.) This solution, of course, applies most clearly to the second courses in a sequence. It probably will not work as well with the third, fourth, or later courses in a sequence if there has been no earlier use of initial enhancement of prerequisites or ML procedures. We hope these ideas will be further explored in the United States as well as in other countries. We believe this solution is relevant at all levels of education, including elementary-secondary, college, and even the graduate and professional school level.

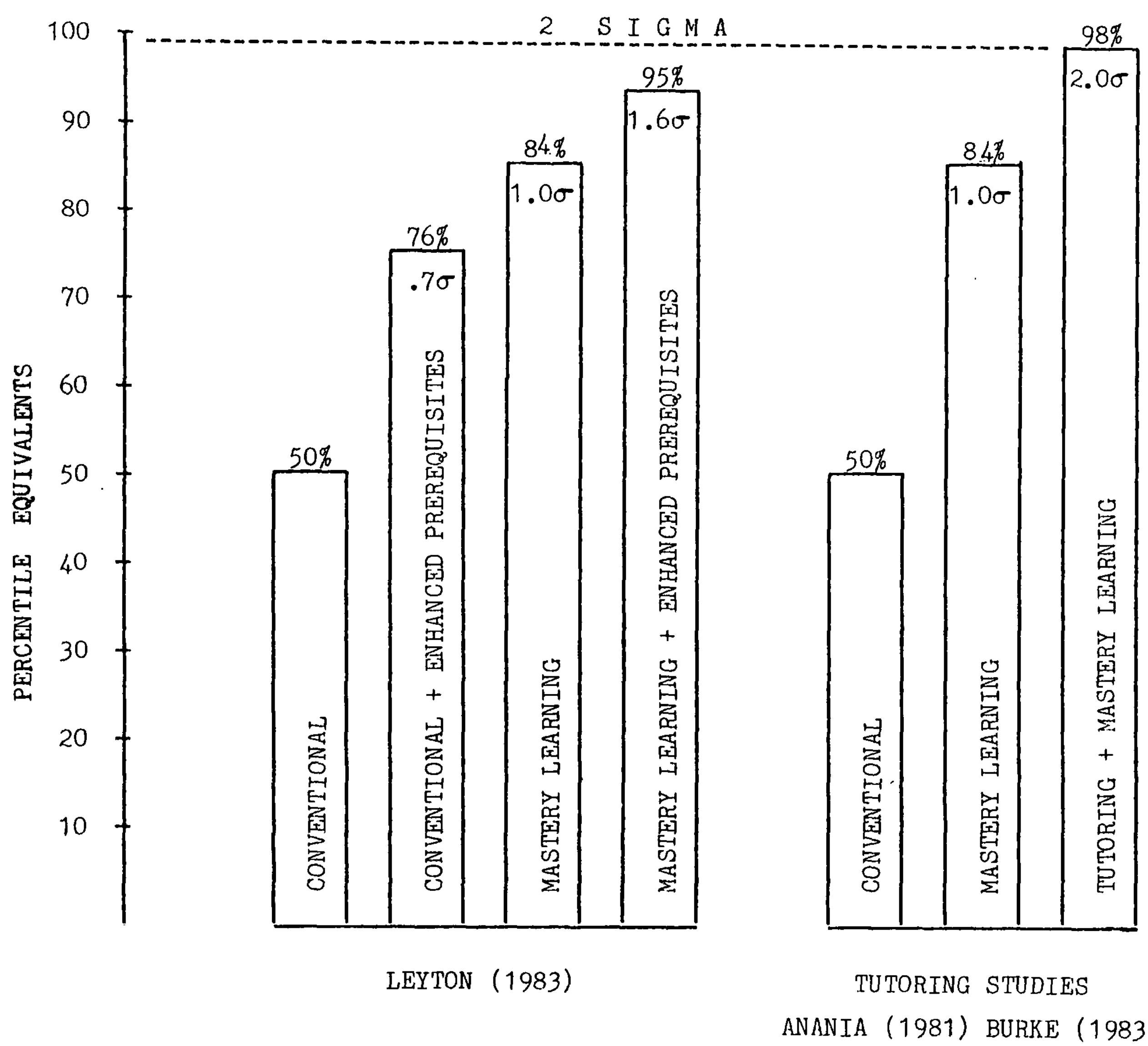
We also regard this approach as widely applicable within a country because the prerequisites for a particular sequential subject or course are likely to be very similar even though different textbooks and teachers may be involved. Thus, a well made test of the initial prerequisites for a particular sequential course—Arithmetic 2, French 2, Reading 2, and so on—may with only minor changes apply to other versions of the same course within a particular country. Also, the procedures that work well in enhancing these prerequisites in one school should work equally well in other schools. Further research is needed to establish the sequential courses in which this approach is most effective.

Finally, the time cost of the initial enhancement procedures is limited to the class hours of the course during the first week of the sequential course, while the time or other costs of the ML procedures have usually been very small. We hope that this approach to the 2 sigma problem will be found to be a widely applicable as well as economical solution available to most teachers who wish to improve student learning, student academic self-concept, and student attitudes and interest in the learning.

Our graduate students have written papers on several other approaches for improving student processing of conventional instruction:

1. Help students develop a student support system in which groups of two or three students study together, help each other when they encounter difficulties in the course, help each other review in advance of taking tests, and review their learning periodically. A student support system that provides support, encouragement and even help when needed can do much to raise the level of learning of the

**FIGURE 2.** *Average summative achievement scores under different learning conditions. Comparison of tutoring studies, mastery learning, and enhanced prerequisites.*



participants. There is evidence that these and other cooperative learning efforts are almost as effective as ML procedures. (Cooperative Learning — Effect size .80 (79 percentile) Slavin, 1980.)

2. There is evidence that students who take special programs to improve their reading and/or their study and learning methods tend to learn more effectively. Ideally, such special programs should be available at the beginning of each new school level, that is, junior high school, high school, and so on. One would hope that the special programs would be closely related to the academic courses the student is currently taking. (Improved reading/study skills—Effect size 1.00 (84 percentile) (Pflaum, Walberg, Karegianes, & Rasher, 1980).

### Improve Instructional Materials and Educational Technology

The textbook in the United States, as well as in most advanced countries in the world, is an almost universal part of school instruction. There has been much work on the improvement of the textbooks for reading and, to some extent, arithmetic, mathematics, and science subjects. Most of these are in relation to special curricular improvements, which include improvements in the sequential nature of the topics, the attempt to find important ideas or schema that help to interrelate the different parts of the subject, and improvements in the illustrations and exercises in the books. However, as far as we can find, these improvements have not had very significant effects on student achievement unless the teachers were provided with much inservice education for the new curriculum or the new textbook.

My graduate students and I have been intrigued by the possibility that the organization of a particular section (or chapter) of the textbook might be better integrated or the parts of the section more closely related to each other. Preorganizers or advanced organizers (Ausubel, 1960), have been moderately effective when provided in the textbook or provided by the teacher at the beginning of the new unit of the course. These may be provided in the form of objectives, some ideas about what will be learned in the unit, or a brief discussion of the relation between what has already been learned and what will be learned in the unit. Such advanced organizers (Luiten, Ames, & Ackerson, 1980) appear to have an average effect size on achievement of about .2 sigma. (Incidentally, such advance organizers have about a .4 sigma effect on retention of the learning.) Although this effect is rather consistent, by itself it is not enough to contribute significantly to the 2 sigma effect. It is likely that a *combination* of advance organizers at the beginning of a new topic, further organizational aids during the chapter or unit, as well as appropriate questions, summaries, or other organizational aids at the end of the unit, may have a substantial effect on the student's learning of that chapter.

*In Process*

One of our students, Carlos Avalos, is working on a study of the effect of *organizational aids* in the instructional material combined with the *initial enhancement of cognitive prerequisites* and the ML *feedback-corrective* procedures. Avalos is planning a research design that will enable him to determine the separate effects of each of the three processes, the effect of any two of the processes, and the combined effect of all three processes. At the least, it is anticipated that the combination of any two of the processes will be greater than the effects of any one of the same processes. It is hoped that the effect of any two will be above 1.3 sigma (90 percentile). If this is found, it will provide several new solutions to the 2 sigma problem—some of which can be done with very little cost or effort by the teachers or the school system.

Avalos expects the results noted above because the organizational aids can be built into new textbooks and can be used by the students with a minimum of emphasis by the teachers. The initial enhancement of the prerequisites is completed before the students begin the study of the new course subject matter, whereas the ML feedback-corrective procedures take place every 2 or 3 weeks during the course. We believe that each of these processes is somewhat independent of the other processes.

Other suggestions for the improvement of instructional materials and educational technology include the following:

1. Some of our students have used computer learning courses, such as the Plato system, which appear to work very well for highly motivated students. We believe that it should be possible to determine whether particular computer courses enable sizeable proportions of students to attain the 2 sigma achievement effect. The effectiveness of the computer courses can be determined in terms of the time required, completion rates, student performance on achievement tests, and student retention of the learned material. It is hoped that the more effective computer courses will also have positive effects on such affective characteristics as academic self-concept, interest in the subject, and desire to learn further with computer learning methods.

2. Although the average effect size for new science and math curricula in the United States is only .3 sigma, some of the new curricula (or textbooks) in these and other subjects may be much more effective than others. We propose a careful search of the new curricula and textbooks to determine which ones are more effective and to determine what characteristics make them more effective than the others.

### Home Environment and the Peer Group

In this section, we are primarily concerned with the out-of-school support that the student receives from the home or the peer group. We are interested in the ways in which the student's achievement, academic aspirations and goals, and progress in learning are influenced by these types of support. We know that the home environment does have great influence on the pupil's school learning and that this influence is especially effective at the elementary school level or earlier. The peer group's influence is likely to be strongest (both positively or

*Although it is difficult to influence the student's choice of friends and peer groups, the availability in the school of a variety of extracurricular activities and clubs... should enable students to be more selective in their peer choices within the school setting.*

negatively) at the secondary school level.

*Home Environment Processes*

There have been a large number of studies of the home environment processes that affect the students' school learning. These studies involve interviews and observations directed at determining the relevant interactions between parents and their children. The studies find correlations of +.70 to +.80[3] between an index of the home environment processes and the children's school achievement. Some of the home environment processes that appear to have high relationships with school achievement include the following:

1. Work habits of the family—the degree of routine in the home management, the emphasis on regularity in the use of space and time, and the priority given to schoolwork over other more pleasurable activities.
2. Academic guidance and support—the availability and quality of the help and encouragement parents give the child for his or her schoolwork and the conditions they provide to support the child's schoolwork.
3. Stimulation in the home—the opportunity provided by the home to explore ideas, events, and the larger environment.
4. Language development—opportunities in the home for the development of correct and effective language usage.
5. Academic aspirations and expectations—the parents' aspirations for the child, the standards they set for the child's school achievement, and their interests in and knowledge of the child's school experiences.

These studies of the home environment processes began with the work of Dave (1963) and Wolf (1964, 1966), and since then have been replicated in other studies done in the United States and other countries (Marjoribanks, 1974; Kalinowski & Sloane, 1981).

These previous studies of the relationship between the home and the children's school achievement suggest a strong effect of the home environment on the school learning of the children, but they do not provide evidence on the extent to which the home environment can be *altered* and the effect of such alteration on changes in the children's school achievement.

A recent study done in Thailand by Janhom (1983) involved a control group and three experimental groups of parents (and their children). In this study, the most effective treatment of the parents was for the group of parents to meet with a parent educator for about 2 hours twice a month for 6 months. In these meetings, the parents discussed ways in which they could support their children's learning in the school. There was usually an initial presentation made by the parent educator on one of the home environment processes and then the parents discussed what they did as well as what they hoped to do to support their children's school learning.

Another experimental approach included visits to each home separately by a parent educator twice a month for 6 months. A third experimental approach was that newsletters about the same topics were sent to the home twice a month for 6 months.

The parents of all four groups were observed and interviewed at the beginning and end of the 6-month period using the Dave (1963) interview and observational methods. Although the three experimental approaches show significantly greater changes in the parents' home environment index than the control group, the most effective method was the series of meetings between groups of parents and the parent educator. The changes in the home environment of this group were highly significant when compared with the changes in the other three groups of parents.

The fourth grade children of all these parents were given a national standardized test on reading and mother tongue as well as arithmetic at the beginning and end of the 6-month period. It was found that the children of the meeting group of parents had changed by 1 sigma in achievement, as contrasted with the change in the control group of children. In comparison, the parent educators' visit to each of the homes every other week had only a .5 sigma effect on the children's school achievement.

Other methods of changing the home environment have been reported by Dolan (1980), Bronfenbrenner (1974), and Kalinowski and Sloane (1981). Again, the most effective approaches to changing the home environment processes result in changes in the children's school achievement. (Home Environment — Effect size .50 (69 percentile), Iverson & Walberg, 1982.)

The methods of changing the home environments are relatively costly in terms of parent educators meeting with groups of parents over a series of semi-monthly meetings, but the payoff of this approach is likely to be very great. If parents continue to encourage and support each of their children to learn well in school throughout the elementary school years, this should greatly help the children during the years they will attend schools and colleges.

Although such research has not been done as yet, we hope that others will explore an approach to the 2 sigma problem of providing effective parent education combined with the mastery learning method. Because parent support takes place in the home and ML takes place in the school, we expect that these two effects will be additive. The result should be close to a 2 sigma improvement in student learning.

Ideally, if both methods began with first or second grade children, one might hope that the combination would result in consistently good learning, at least through the elementary school years, with less and less need for effort expended by the parents or by the use of ML procedures in the school.

*Peer Group*

During the adolescent years, it is likely that the peer group will have considerable influence on the student's activities, behavior, attitudes, and academic expectations. The peer group(s) to which the individual "belongs" also has some effect on the student's high school achievement level as well as further academic aspirations. These effects appear to be greatest in urban settings. Although it is difficult to influence the student's choice of friends and peer groups, the availability in the school of a variety of extracurricular activities and clubs (e.g., athletics, music, science, mathematics, social, etc.) should enable students to be more selective in their peer choices within the school setting. (Peer Group Influence—Effect size .20 (58 percentile) (Ide, Haertel, Parkerson, & Walberg, 1981).

### Improvement of Teaching

When we compare student learning under conventional instruction and tutoring we note that approximately 20% of the students under conventional instruction do about as well as the tutored students. (See Figure 1). That is, tutoring probably would not enable these top students to do any better than they already do under conventional instruction. In contrast, about 80% of the students do relatively poorly under conventional instruction as compared with what they might do under tutoring. We have pondered these facts and believe that this in part results from the unequal treatment of students within most classrooms.

Observations of teacher interaction with students in the classroom reveal that teachers frequently direct their teaching and explanations to some students and ignore others. They give much positive reinforcement and encouragement to some students but not to others, and they encourage active participation in the classroom from some students and discourage it from others. The studies find that typically teachers give students in the top third of the class the greatest attention and students in the bottom third of the class receive the least attention and support. These differences in the interaction between teachers and students provide some students with much greater opportunity and encouragement for learning than is provided for other students in the same classroom (Brophy & Good, 1970.)

It is very different in a one-to-one tutoring situation where there is a constant feedback and corrective process between the tutor and the tutee. If the explanation is not understood by the tutee, the tutor soon becomes aware of it and explains it further. There is much reinforcement and encouragement in the tutoring situation, and the tutee must be actively participating in the learning if the tutoring process is to continue. In contrast, there is less feedback from each student in the group situation to the teacher—and frequently the teacher gets most of the feedback on the clarity of his or her explanations, the effect of the reinforcements, and the degree of active involvement in the learning from a *small* number of high achieving students in the typical class of 30 students.

Teachers are frequently unaware of the fact that they are providing more favorable conditions of learning for some students than they are for other students. Generally, they are under the impression that all students in their classes are given equality of opportunity for learning. One basic assumption of our work on teaching is the belief that when teachers are helped to secure a more accurate picture of their own teaching methods and styles of interaction with their students, they will increasingly be able to provide more favorable learning conditions for more of their students, rather than just for the top fraction of the class.

In some of our research on the 2 sigma problem, we have viewed the task of teaching as providing for more equal treatment of students. We have been trying to give teachers feedback on their differential treatment of students. We attempt to provide teachers with a mirror of what they are now doing and have them develop techniques for equalizing their interactions with the students. These include such techniques as: (a) attempt to find something positive and encouraging in each student's response, (b) find ways of involving more of the students in active engagement in the learning process, (c) secure feedback from a small random sam-

*Teachers are frequently unaware of the fact that they are providing more favorable conditions of learning for some students than they are for other students. Generally, they are under the impression that all students in their classes are given equality of opportunity for learning.*

ple of students to determine when they comprehend the explanations and illustrations, and (d) find ways of supplying additional clarification and illustrations as needed. The major emphasis in this work was *not* to change the teachers' *methods* of instruction, but to have the teacher become more aware of the ways in which he or she could more directly teach to a cross section of the students at each class section.

The first of our studies on improving instruction was done by Nordin (1979, 1980), who found ways of improving the cues and explanations for students as well as increasing the active participation of students.

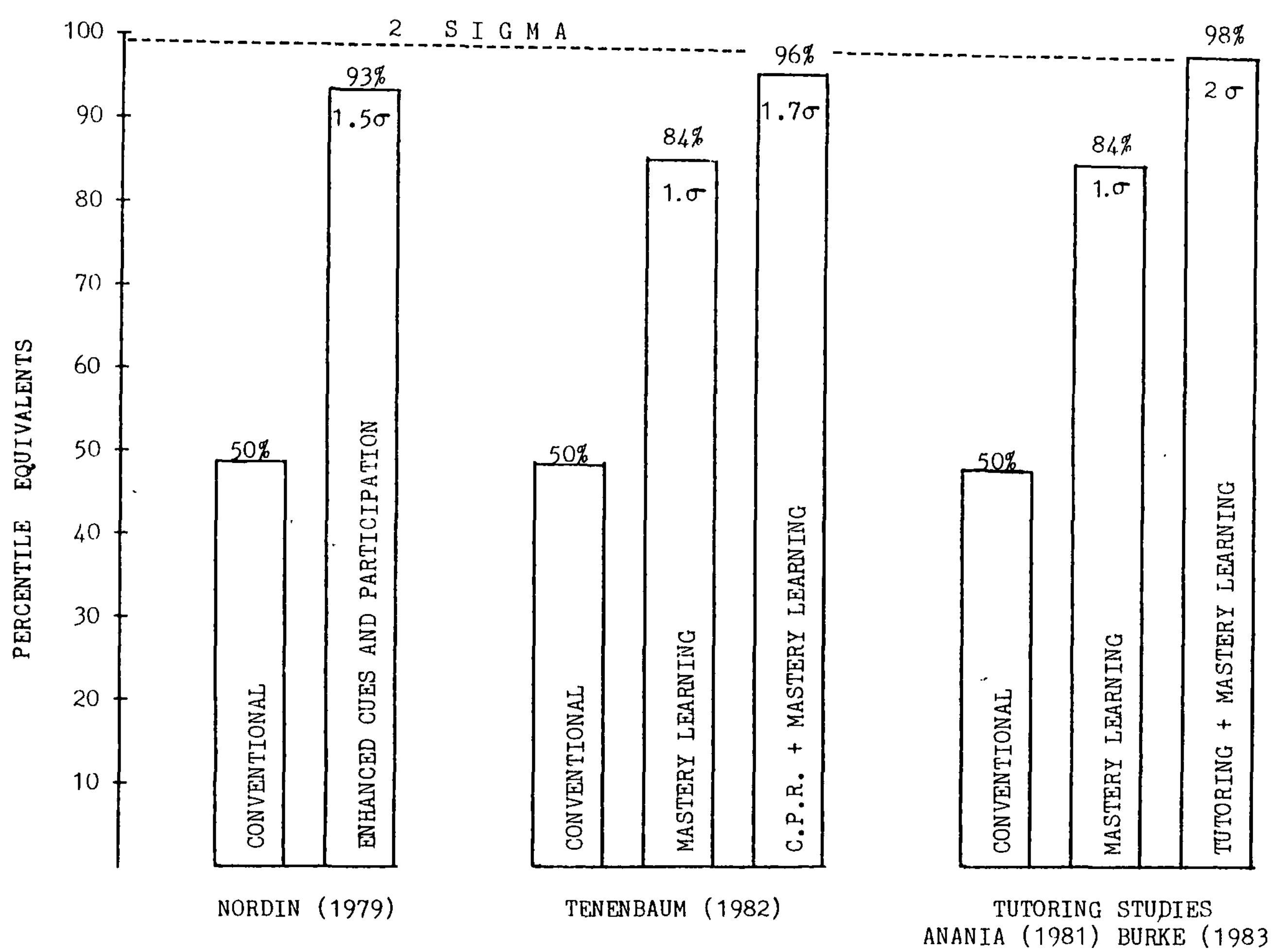
He found it helpful to meet frequently with the teachers to explain these ideas as well as to observe the teachers and help them determine when they still needed to improve these qualities of the instruction. He also had independent observers noting the frequency with which the experimental teachers were using these ideas well or poorly. Similarly, he had students note the frequency with which they were actively participating in the learning and any problems they had with understanding the ideas or explanations.

In this research he compared student learning under conventional instruction and under enhanced cues (explanations) and participation conditions. During the experiment, observers noted that the student participation and the explanations and directions were positive in about 57% of the observations in the control class as compared with about 67% in the enhanced cue + participation classes. Students in the control classes noted that the cues and participation were positive for them about 50% of the time as compared with about 80% of the time for the students in the enhanced cue + participation classes.

In terms of final achievement, the average student in the enhanced cue and participation group was 1.5 sigma higher than the average student in the control classes. (The average student in the enhanced group was above 93% of the students in the control classes.) (See Figure 3.) Nordin (1979, 1980) also made use of the ML procedures in other classes and found that they worked even better than the enhanced cue + participation procedures. Unfortunately, he did not use the ML in combination with the enhanced cue + participation methods.

In any case, Nordin (1979, 1980) did demonstrate that teachers could be taught ways to be more responsive to most of the students in the class, secure increased participation of the students, and insure that

**FIGURE 3.** *Average summative achievement scores under different learning conditions. Comparison of tutoring studies, mastery learning, and enhanced instructional methods.*



most of the students understood the explanations and illustrations that the teacher provided. The observers noted that the students in the enhanced participation and cue classes were actively engaged in learning (time on task) about 75% of the classroom time, whereas the control students were actively learning only about 57% of the time.

In a later study, Tenenbaum (1982) compared control groups, ML groups, and Enhanced Cues, Participation, and Reinforcement in combination with ML (CPR + ML). Tenenbaum studied these three methods of teaching with randomly assigned students in two different courses—sixth grade science and ninth grade algebra.

Tenenbaum also used student observation of their own classroom processes on cues, participation, and reinforcement. He found that under the CPR + ML, students responded positively about their own participation about 87% of the time as contrasted with 68% in the control classes.

The results of this study demonstrated large differences between the three methods of instruction with the final achievement scores of the CPR + ML group about 1.7 sigmas above the control students (the average student in this group was above 96% of the students in the control group.) The average student in the ML groups was the usual 1 sigma above the control students. (See Figure 3).

We believe that this research makes it clear that teachers in both the Nordin and Tenenbaum studies could (at least temporarily) change their teaching methods to provide more equal treatment of the students in their classes. When this more equal treatment is provided and supplemented with the ML feedback and corrective procedures, the average student approaches the level of learning found under tutoring methods of instruction.

We believe there are a variety of methods of giving feedback to teachers on the extent to which they are providing equality of interaction with their students. The tactic of providing a "mirror" to the teacher of the ways in which he or she is providing cues and explanations, appropriate reinforcement, and securing overt as well as covert participation of the students in the learning, seems to us to be an excellent approach. This may be in the form of an observer's notes on what the teacher and students did, student observations of their own interactions with the teaching (preferably anonymous, but coded as to whether the students are in the top third, middle third, or the bottom third of the class in achievement), such as their understanding of the cues and explanations, the extent of their overt & covert participation, and the amount of reinforcement they are getting. Perhaps a video-tape or audio tape recording of the class could serve the same purpose if the teacher is given brief training on ways of summarizing the classroom interaction between the teacher and the students in the class.

It is our hope that when teachers are helped to secure a more accurate picture of their own teaching methods and styles of interaction with their students, they will be better able to provide favorable learning conditions for most of their students.

**Improvement of Teaching of the Higher Mental Processes**

Although there is much of rote learning in schools through the world, in some of the national curriculum centers in different countries (e.g., Israel, Malaysia, South Korea) I find great emphasis on problem-solving, application of principles, analytical skills, and creativity. Such higher mental processes are emphasized because these centers believe that they enable the student to relate his or her learning to the many problems he or she encounters in day-to-day living. These abilities are also stressed because they are retained and used long after the individual has forgotten the detailed specifics of the subject matter taught in the schools. These abilities are regarded as one set of essential characteristics needed to continue learning and to cope with a rapidly changing world. Some curriculum centers believe that these higher mental processes are important because they make learning exciting and constantly new and playful.

In these countries, subjects are taught as methods of inquiry into the nature of science, mathematics, the arts, and the social studies. The subjects are taught as much for the ways of thinking they represent as for their traditional content. Much of this learning makes use of observations, reflections on these observations, experimentation with phenomena, and the use of first hand data and daily experiences, as well as the use of primary printed sources. All of this is reflected in the materials of instruction, the learning and teaching processes used, and the questions and problems used in the quizzes and formative testing, as well as on the final summative examinations.

In sharp contrast with some of these other countries, teachers in the United States typically make use of textbooks that rarely pose real problems. These textbooks emphasize specific content to be remembered and give students little opportunity to discover underlying concepts and principles and even less opportunity to attack real problems in the environments in which they live. The teacher-made tests (and standardized tests) are largely tests of remembered information. After the sale of over one million copies of the *Taxonomy of Educational Objectives—Cognitive Domain* (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956) and over a quarter of a century of use of this domain in preservice and in-service teacher training, it is estimated that over 90% of test questions that U.S. public school students are *now* expected to answer deal with little more than information. Our instructional material, our classroom teaching methods, and our testing methods rarely rise above the lowest category of the Taxonomy-knowledge.

In the tutoring studies reported at the beginning of this paper, it was found that the tutored students' Higher Mental Process (HMP) achievement was 2.0 sigma above the control students. (See Figure 4.) (The average tutored student was above 98% of the control students on the HMP part of the summative examination.) It should be noted that in these studies higher mental processes as well as lower

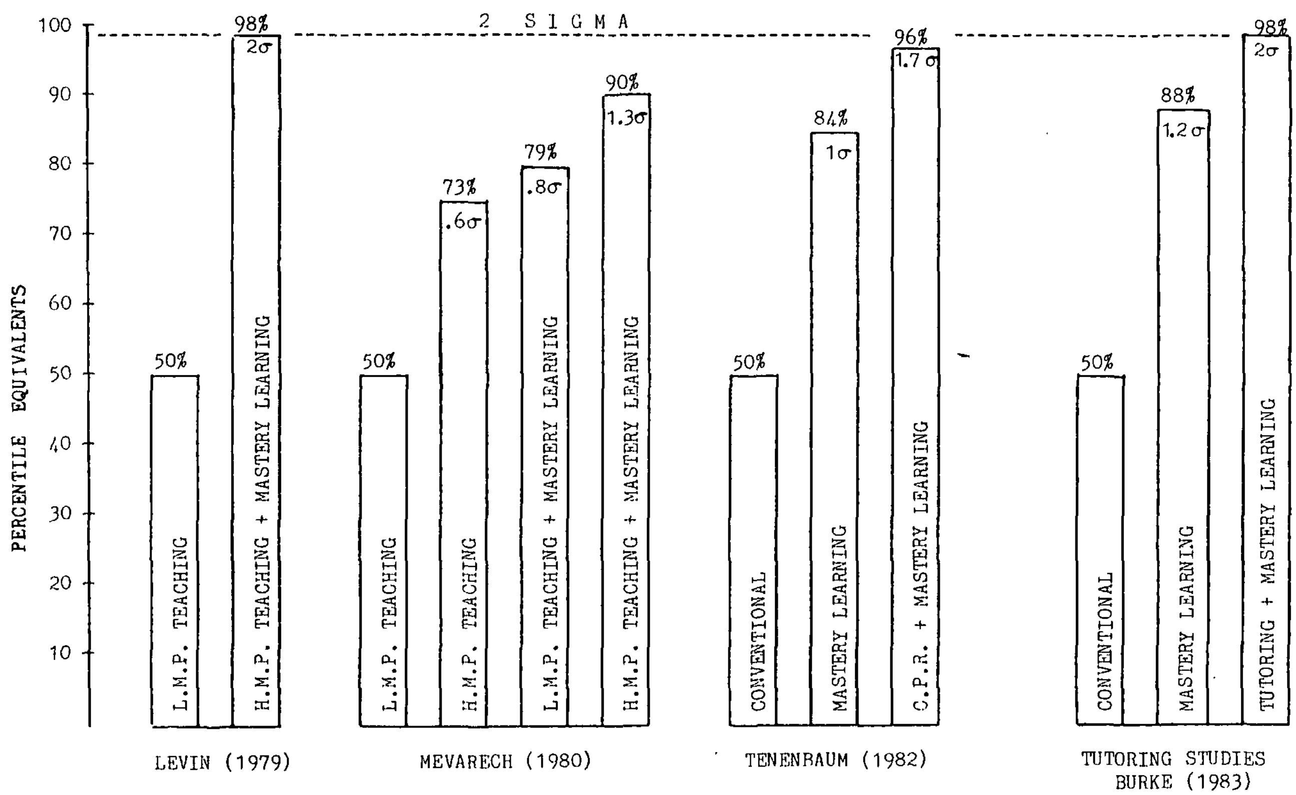



**FIGURE 4.** *Average higher mental process achievement under different learning conditions. Comparison of tutoring studies, mastery learning, and higher mental process instructional methods.*

mental process questions were included in the formative tests used in the feedback-corrective processes for both the ML and tutored groups. Again, the point is that students can learn the higher mental processes if they become more central in the teaching-learning process.

Several studies have been made in which the researcher was seeking to improve the higher mental processes.

We have already referred to the Tenenbaum (1982) study, which emphasized changing teacher-student interaction. In this study, the Cue-Participation-Reinforcement + Mastery Learning student group was 1.7 sigma higher than the control students on the higher mental process part of the summative examination. (The average CPR + ML student was above 96% of the control students on the higher mental processes.) (See Figure 4.)

Another study done by Levin (1979) was directed to improving the higher mental processes by emphasizing the mastery of the lower mental processes and providing learning experiences in which the students applied principles in a variety of different problem situations. On the summative examinations, the students were very high on the knowledge of principles and facts and in their ability to apply the principles in new problem situations. These experimental students were compared with a control group that was only taught the principles (but not their application). On the higher mental processes, the experimental group was 2 sigma above the control students (the average experimental student was above 98% of the control students) in the ability to apply the principles to new problem situations.

A third study by Mevarech (1980) was directed at improving the higher mental processes by emphasizing heuristic problem solving and including higher and lower mental process questions in the formative testing and in the feedback-corrective processes. On the higher mental process part of the summative tests, the group using the heuristic methods + ML (HMP Teaching + ML) was 1.3 sigma above the control group (L.M.P. Teaching) taught primarily by learning algorithms—a set of rules and procedures for solving particular math problems (the average student in this experimental group was above 90% of the control students).

In all of these studies, attempts to improve higher mental processes included group instruction emphasizing higher mental processes and feedback-corrective processes, which also emphasized higher mental processes. In addition, the tutoring studies included an instructional emphasis on both higher and lower mental processes, as well as the feedback-corrective processes, which included both higher and lower mental processes. It was evident in all of these studies that in the formative feedback and corrective processes the students needed and received more corrective help on the higher mental processes questions and problems than they did on the lower mental process questions.

## Summary

The Anania (1982, 1983) and Burke (1984) studies comparing student learning under one-to-one tutoring, ML, and conventional group instruction began in 1980. As the results of these separate studies at different grade levels and in different school subjects began to emerge, we were astonished at the consistency of the findings as well as the great differences in student cognitive achievement, attitudes, and academic self-concept under tutoring as compared with the group methods of instruction.

During the past 4 years, the graduate students in my seminars at the University of Chicago and Northwestern University considered various approaches to the search for group methods of instruction that might be as effective as one-to-one tutoring. This paper reports on the research studies these students have completed, the studies that are still in process, and some of the other ideas we explored in these seminars.

Although all of us at first thought it was an impossible task, we did agree that if we succeeded in finding *one* solution, there would soon be a great many solutions. In this paper, I report on six solutions to the 2 sigma problem. In spite of the difficulties, our graduate students found the problem to be very intriguing because the goal was so clear and specific—*find methods of group instruction as effective as one-to-one tutoring.*

Early in the work, it became evident that more than group instruction in the school had to be considered. We also needed to find ways of improving the students' learning processes, the curriculum and instructional materials, as well as the home environmental support of the students' school learning. This paper is only a preliminary report on what has been accomplished to date, but it should be evident that much can now be done to improve student learning in the schools. However, the search is far from complete. We look for additional solutions to the 2 sigma problem to be reported in the next few years. I hope some of the readers of this article will also find this problem challenging.

## APPENDIX

*Effect Size References*


**Tutorial Instruction***

Anania, J. (1982). The effects of quality of instruction on the cognitive and affective learning of students (Doctoral dissertation, University of Chicago, 1981). *Dissertation Abstracts International, 42,* 4269A.

Burke, A.J. (1984). Students' potential for learning contrasted under tutorial and group approaches to instruction (Doctoral dissertation, University of Chicago, 1983). *Dissertation Abstracts International, 44,* 2025A.

**Reinforcement**

Lysakowski, R.S., & Walberg, H.J. (1981). Classroom reinforcement: A quantitative synthesis. *Journal of Educational Research, 75,* 69–77.

**Feedback-Corrective, Cues & Explanations, and Student Classroom Participation**

Lysakowski, R.S., & Walberg, H.J. (1982). Instructional effects of cues, participation, and corrective feedback: A quantitative synthesis. *American Educational Research Journal, 19,* 559–578.

**Student Time on Task (in the classroom)**

Frederick, W.C., & Walberg, H.J. (1980). Learning as a function of time. *Journal of Educational Research, 73,* 183–194.

**Improved Reading/Study Skills**

Pflaum, S.W., Walberg, H.J., Karegianes, M.L., & Rasher, S. (1980). Reading instruction: A quantitative synthesis. *Educational Researcher, 9,* 12–18.

**Cooperative Learning**

Slavin, R.E. (1980). Cooperative learning. *Review of Educational Research, 50,* 315–342.

**Home Work (graded) and Home Work (assigned)**

Paschal, R., Weinstein, T., & Walberg, H.J. (in press). Effects of homework: A quantitative synthesis. *Journal of Educational Research.*

**Classroom Morale**

Haertel, G.D., Walberg, H.J., & Haertel, E.H. (1981). Social-psychological environments and learning: A quantitative synthesis. *British Educational Research Journal, 7,* 27–36.

**Initial Cognitive Prerequisites***

Leyton, F.S. (1983). The extent to which group instruction supplemented by mastery of the initial cognitive prerequisites approximates the learning effectiveness of one-to-one tutorial instruction (Doctoral dissertation, University of Chicago, 1983). *Dissertation Abstracts International, 44,* 974A.

**Home Environment Intervention (parental educational program)**

Iverson, B.K., & Walberg, H.J. (1982). Home environment and learning: A quantitative synthesis. *Journal of Experimental Education, 50,* 144–151.

**Peer & Cross-Age Remedial Tutoring**

Cohen, P.A., Kulik, J.A., & Kulik, C.C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal, 19,* 237–248.

**Higher Order Questions**

Redfield, D.L., & Rousseau, E.W. (1981). Meta-analysis of experimental research on teacher questioning behavior. *Review of Educational Research, 51,* 235–245.

**New Science & Math Curricula and Teacher Expectancy**

Walberg, H.J. (1984). Improving the productivity of America's schools. *Educational Leadership, 41,* 8, 19–27.

**Peer Group Influence**

Ide, J., Haertel, G.D., Parkerson, J.A., & Walberg, H.J. (1981). Peer-group influences on learning: A quantitative synthesis. *Journal of Educational Psychology, 73,* 472–484.

**Advance Organizers**

Luiten, J., Ames, W., & Ackerson, G. (1980). A meta-analysis of the effects of advance organizers on learning and retention. *American Educational Research Journal, 17,* 211–218.

*not effect size studies

## Notes

[1]In giving the percentile equivalent we make use of the normal curve distribution. The control class distributions were approximately normal, although the mastery learning and tutoring groups were highly skewed.

[2]$\frac{\text{Mean experimental—Mean control}}{\text{standard deviation of the control}} =$

$$\frac{M_{ex}-M_c}{\text{sigma of control}} = \textit{effect size.}$$

[3]When questionnaires rather than interviews and observations have been used, the correlations are somewhat lower, with the average being between +.45 and +.55.

## References


Anania, J. (1982). The effects of quality of instruction on the cognitive and affective learning of students. (Doctoral dissertation, University of Chicago, 1981). *Dissertation Abstracts International, 42*, 4269A.

Anania, J. (1983). "The influence of instructional conditions on student learning and achievement." *Evaluation in Education: An International Review Series, 7*, 1, 1-92.

Ausubel, D. (1960). The use of advanced organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology, 51*, 267-272.

Bloom, B.S. (1980). The new direction in educational research: alterable variables. *Phi Delta Kappan, 61*, 6, 382-385.

Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H., & Krathwohl, D.R. (1956). *Taxonomy of Educational Objectives: Handbook I, Cognitive Domain.* New York: Longman.

Bronfenbrenner, U. (1974). Is Early Intervention Effective? In H.J. Leichter, (Ed.), *The Family as Educator.* New York: Teachers College Press.

Brophy, J.E., & Good, T.L. (1970). Teachers' communication of differential expectations for children's classroom performance: Some behavioral data. *Journal of Educational Psychology, 61*, 365-374.

Burke, A.J. (1984). Students' potential for learning contrasted under tutorial and group approaches to instruction. (Doctoral dissertation, University of Chicago, 1983). *Dissertation Abstracts International, 44*, 2025A.

Dave, R.H. (1963). The identification and measurement of environment process variables that are related to educational achievement. (Unpublished doctoral dissertation, University of Chicago).

Dolan, L.J. (1980). The affective correlates of home concern and support, instructional quality, and achievement. (Unpublished doctoral dissertation, University of Chicago).

Ide, J., Haertel, G.D., Parkerson, J.A., & Walberg, H.J. (1981). Peer group influences on learning: A quantitative synthesis. *Journal of Educational Psychology, 73*, 472-484.

Iverson, B.K., & Walberg, H.J. (1982). Home environment and learning: A quantitative synthesis. *Journal of Experimental Education, 50*, 144-151.

Janhom, S. (1983). Educating parents to educate their children. (Unpublished doctoral dissertation, University of Chicago).

Kalinowski, A., & Sloane, K. (1981). The home environment and school achievement. *Studies in Educational Evaluation, 7*, 85-96.

Levin, T. (1979). Instruction which enables students to develop higher mental processes. *Evaluation in Education: An International Review Series, 3*, 3, 173-220.

Leyton, F.S. (1983). The extent to which group instruction supplemented by mastery of the initial cognitive prerequisites approximates the learning effectiveness of one-to-one tutorial methods. (Doctoral dissertation, University of Chicago, 1983). *Dissertation Abstracts International, 44*, 974A.

Luiten, J., Ames, W., & Ackerson, G. (1980). A meta-analysis of the effects of advance organizers on learning and retention. *American Educational Research Journal, 17*, 211-218.

Marjoribanks, K. (1974). *Environments for Learning.* London: National Foundation for Educational Research.

Mevarech, Z.R. (1980). The role of teaching-learning strategies and feedback-corrective procedures in developing higher cognitive achievement. (Unpublished doctoral dissertation, University of Chicago).

Nordin, A.B. (1979). The effects of different qualities of instruction on selected cognitive, affective, and time variables. (Unpublished doctoral dissertation, University of Chicago).

Nordin, A.B. (1980). Improving learning: An experiment in rural primary schools in Malaysia. *Evaluation in Education: An International Review Series, 4*, 2, 143-263.

Pflaum, S.W., Walberg, H.J., Karegianes, M.L., & Rasher, S. (1980). Reading instruction: A quantitative synthesis. *Educational Researcher, 9*, 12-18.

Slavin, R.E. (1980). Cooperative Learning. *Review of Educational Research, 50*, 315-342.

Tenenbaum, G. (1982). A method of group instruction which is as effective as one-to-one tutorial instruction. (Doctoral dissertation, University of Chicago, 1982). *Dissertation Abstracts International, 43*, 1822A.

Walberg, H.J. (1984). Improving the productivity of America's schools. *Educational Leadership, 41*, 8, 19-27.

Wolf, R.M. (1964). The identification and measurement of home environmental process variables that are related to intelligence. (Unpublished doctoral dissertation, University of Chicago).

Wolf, R.M. (1966). The Measurement of Environments. In A. Anastasi (Ed.), *Testing Problems in Perspective.* Washington D.C.: American Council on Education.